

To build bridges, or to burn them

Environmentalists who have grown impatient with science and technology need not be dismissed as beyond the reach of reason.

Not everyone's opinion is equally valuable. If this statement startles, then that is testament to a culture in which everyone's say-so is deemed to be of value, whether it emerges in the classroom, on talk radio, in a blog or in political activism. It would be a stretch to allege that this culture created the small group of environmental activists whose tactics include the arson of scientific laboratories. But a case can certainly be made that the two are connected.

In this issue, *Nature* tells the story of a radical environmental movement in the northwestern United States that allegedly mounted physical attacks on the laboratories of scientists whose investigations they deemed undesirable (see page 498). The numbers of those actually involved in the attacks was rather low, but their base of support is considerably larger.

Environmentalist scientists should maintain a dialogue with this larger group. But this will be no easy task: signs of paternalism or scepticism about emotional arguments will quickly alienate a section of public opinion whose views, although logically fuzzy, are very firmly held.

The perpetrators of the attacks in question identified themselves as members of the Earth Liberation Front (ELF). They set fires at labs that they believed to be engaged in the genetic engineering of trees, for example, or in the culling of wildlife. Their motivation was doubtless provided by a conviction that their own judgement about the morality of the science in question trumped everyone else's — supplemented, perhaps, by the thrill of the criminal action itself.

Emotional response

ELF communiqués and writings, as well as interviews with supporters of these crimes, show clearly that many such activists have moved in circles in which intuition, emotional response and individualism are pre-eminent values. Such activists can be seen as an American species of Romantic — except rather than merely contemplating the sublime, they feel inclined to mount attacks in its defence.

It is all too easy to imagine how this outlook could come into conflict with the scientific method, which rejects emotion in favour of objective measurement and rationality, and which moves forward by subjecting itself to critical inquiry.

But this inherent conflict is greatly inflamed and inflated by the view that many members of the activist communities in question take of scientists: as an oppressive hierarchy of élites making pronouncements dressed up as facts. Many of them feel that these pronouncements are readily corruptible by the interests of the rich and powerful. Some feel that what were touted as scientific solutions to society's problems — such as DDT, thalidomide and nuclear power — turned out not to deliver what they promised. So many of this new generation of environmental activists hold science and technology in deep mistrust.

The small minority of that generation who are responsible for physical attacks on laboratories and other targets possess youth, passion and uncheckable self-assurance. They are angry, idealistic, and have a sense of place and a sense of urgency. They see themselves as fighting for a just cause, and in its pursuit they don't need science any more than do the creationists, with whom they have little else in common. Their hearts tell them the truth. And, in an individualistic, relativistic culture, they see their opinions as just as valid as anyone else's. They don't approve of some lines of inquiry, especially anything to do with genetic engineering or testing on animals. And so: boom.

Several supporters of these drastic actions say that they have nothing but disdain for long, patient field studies, computer models or the careful exploration of data. They say that the environment is in crisis, and that the time for action is now. Some of them think that science itself is not worth keeping — it is a facet, they consider, of a civilization rotten to the core.

Mutual affirmation

Why should scientists bother to reach out to those whose community norms, reinforced by years of enthusiastic, mutual affirmation, run so directly counter to the scientific method? They should do so because this group carries influence within a larger environmental movement that, in turn, constitutes a powerful and potentially constructive element in today's political landscape.

The scientific community's best hope for retaining influence over such difficult territory rests in presenting itself diligently and relentlessly — in schools, in universities, on television and on the Internet — as what it is. That is, not as a set of rules laid down from above by an élite, but as a set of methods for the investigation of the plethora of difficult or near-intractable problems confronting humanity.

The community and its supporters should stress how science can, for example, help us to learn how a natural habitat works, so that its essential elements can be carefully nurtured. Science can also sometimes, if presented with enough aplomb and footnotes, impress policy-makers. And some technology need not consume valuable resources: much of it will save them instead.

The radical radicals, the adherents of the Earth and Animal Liberation Fronts, will never be convinced. But there is also a much larger group of less intense sympathizers who feel strongly about the environment, attend a march here and there, and share the same views on science. They might be more likely to change their minds about science if its practitioners would desist from sneering at emotional argument and demonstrate that science is a window through which we can see our world more clearly. ■

"Many activists feel that scientists' pronouncements are readily corruptible by the interests of the rich and powerful."



Power and particles

String theories dominate for good reason.

For decades, high-energy physicists have basked in the glow of a theory that can apparently do no wrong. Prosaic though the 'standard model' may sound, every experimental observation from accelerators has fulfilled its predictions, often to extraordinary precision. It is also falsifiable, in that it predicts particles that should be observed in the new experimental regimes to be explored by the Large Hadron Collider at CERN from 2007 or 2008 onward. And it's conceptually sweet, exploiting mathematical symmetries to integrate three of the fundamental forces — electromagnetism and the strong and weak nuclear forces — within a single framework.

It has its weaknesses and embarrassments, however. It doesn't predict particle masses: these and many other observed parameters have to be plugged in by hand. Also, it has infinities, due to the fact that its fundamental particles are point-like. A technique ('renormalization') that gets around these infinities was a great piece of inventiveness and enables the model to deliver its sweeping successes, yet also seems like a sleight-of-hand.

Now imagine a theory that incorporates the standard model within it, that includes just one arbitrary parameter rather than many, incorporates the fourth fundamental force — gravity — within its framework, removes the need for renormalization, allows us to describe the extreme conditions of the earliest moments of the Big Bang and also resolves long-frustrating mismatches between quantum mechanics and relativity. Imagine also that such a model requires new types of mathematics in order to make progress, and that work on it repeatedly reveals levels of order that had not previously been appreciated or even suspected.

There is no theory exactly like the above, but string theory is the closest we have to it. The fundamental entities are no longer point-like but have a finite extent. The vibrations of these features of space-time should in principle give rise to observed particles. Even the current embryonic theories suggest observable characteristics — for

example, the liquid rather than gaseous nature of the quark–gluon 'porridge' produced in collisions of gold nuclei. They also suggest enticing insights into controversies surrounding the thermodynamics of black holes.

These are no more than suggestions, and major new uncertainties have opened up alongside new opportunities. There seems to be a vast number of possible universes in such theories. Such embarrassments could prove fatal. But pursuing them is just as likely to lead to insights that make the theory seem all the more inevitable.

Two recent books have attacked the dominance of string theory among the high-energy theoretical community (see pages 491 and 507). They allege that string theory should have proved itself by now and should have made testable predictions. They also claim that string theorists have exerted an intellectual hegemony, accompanied by downright arrogance, over the field that has discouraged alternatives from being pursued.

The complaints of undue influence and arrogance are not without foundation. And string theory is very far from experiment. There is no point in making predictions prematurely, but some theorists seem to positively revel in the lack of them. A prime goal must be to turn this project, ultimately, into a testable science.

This remoteness from experiment reflects the magnitude of the task. It doesn't justify any suggestion that string theory is played out. The many theorists excited by string theory (perhaps ten times as many as those excited by any competing idea, to judge by the meetings) are going where they think the most innovative and intriguing prospects are to be found.

What drives them? There may be esoteric beauties perceivable to some string *cognoscenti*. But much more compelling is the sheer scope of what string theories seem to encompass. Such power, as expressed in the community, has attracted resentment and embitterment. But this power offers insight of unmatched depth and breadth, and the development of its mathematical foundations has been full of tantalizing incident.

Critical-mindedness is integral to all scientific endeavour, but the pursuit of string power deserves undaunted encouragement. ■

One small step

Nature Nanotechnology will spearhead rapid progress in understanding the nanoscale.

Over the past 20 years or so, *Nature* has published several landmark papers in nanotechnology, including the discoveries of C_{60} , also known as buckyballs (H. W. Kroto *et al.* *Nature* 318, 162–163; 1985), and of carbon nanotubes (S. Iijima *Nature* 354, 56–58; 1991). But the sheer volume of papers published in this sphere is rapidly expanding, and the flow of exciting science already far exceeds our capacity to publish it in this general scientific journal. Hence this month's launch of *Nature Nanotechnology*.

The papers in its first issue reflect the breadth of the field, with contributions as diverse as the first demonstration of a carbon-nanotube superconducting quantum-interference device and the development of a memory device based on a virus. The multidisciplinary nature of

this work is, of course, one of the hallmarks of the discipline.

Research agencies and industrial companies around the world are now investing heavily in nanotechnology, in anticipation not just of scientific results but also of a substantial economic return. In electronics, for example, a big reduction in the scale of semiconductor circuitry would permit techniques that operate only at the nanoscale, such as molecular computing. Other applications are expected in manufacturing, materials, energy and environmental technology. There are also tremendous opportunities in medicine — particularly in imaging and drug delivery.

Past experience suggests that the launch of this new research journal will strengthen, rather than weaken, *Nature* itself, which will continue to publish papers in this field. *Nature Nanotechnology* will benefit authors and readers by providing greater exposure for the topic, as well as healthy competition for established journals. We are confident that it will play a central role in the probing of this valuable sphere of knowledge. ■

RESEARCH HIGHLIGHTS

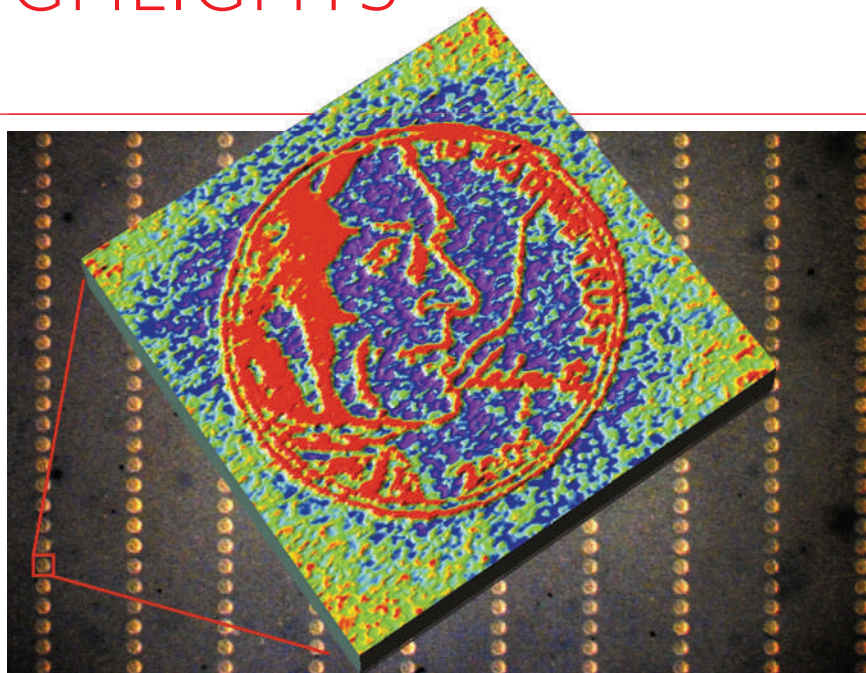
The big draw

Angew. Chem. Int. Ed. **45**, 1–4 (2006)

Fifty-five thousand pens, all writing at once, have been used to draw the same number of images of Thomas Jefferson in a square centimetre. Each portrait (pictured) is drawn with lines of dots 80 nanometres wide.

The 'pens' are silicon nitride tips attached to cantilevers and fixed side by side into a vast array by Chad Mirkin of Northwestern University in Evanston, Illinois, and his colleagues.

They used this equipment to perform massively parallel dip-pen nanolithography, in which 'ink' on the tips is transferred to a substrate at nanoscale resolution. This has a number of possible applications, and its use as a tool to build protein arrays for biological research is being explored.



ANGEWANDTE CHEMIE

CANCER BIOLOGY**A short-lived recovery**

J. Clin. Invest. **116**, 2610–2621 (2006)

Some promising cancer therapies work by starving tumours of their blood supply, but what happens after treatment stops?

Donald McDonald of the University of California, San Francisco, and his colleagues treated tumours in mice with drugs that inhibit signalling by VEGF, a protein that promotes blood-vessel growth. Tumour blood vessels regressed, leaving behind an empty layer of connective tissue known as the basement membrane.

Unfortunately, only a week after therapy ceased, the tumours had been fully repopulated by active blood vessels. The authors speculate that the residual basement membrane acts as a scaffold, with new vessels growing through the tracks of the old ones. This tissue could be a worthwhile target for future therapies.

ECOLOGY**Flower arrangement**

J. Evol. Biol. doi:10.1111/j.1420-9101.2006.01216.x (2006)

A project in California has unearthed hints that pollinators had a role in the divergence of the local monkey flower (*Mimulus aurantiacus*), which has red blooms near the coast and yellow blooms inland.

Evolutionary biologists Matthew Streisfeld and Joshua Kohn at the University of California, San Diego, studied the habits of hummingbirds and hawkmoths, which pollinate the flowers. They found that

hummingbirds strongly preferred red flowers to yellow ones, whereas hawkmoths visited yellow flowers more than 99% of the time.

The researchers say that the inland range of red flowers may increase in future, because the hummingbird population, able to forage year-round in gardens, is thriving.

DRUG DISCOVERY**Pattern recognition**

Science **313**, 1929–1935 (2006); *Cancer Cell* doi:10.1016/j.ccr.2006.09.005 and 10.1016/j.ccr.2006.09.006 (2006)

Researchers have compiled a database of drug-associated gene-expression profiles to help in the search for new drugs.

Todd Golub of the Broad Institute in Cambridge, Massachusetts, and his colleagues used DNA microarrays, which reveal the level of activity of every known gene in a particular tissue, to measure the effects of 164 compounds on human cancer cells. Having input their results

into a database, they designed software to make it possible to take any new microarray pattern and search the database for similar or opposite patterns.

This concept, reported in *Science*, has already proved its worth. Scott Armstrong at Children's Hospital Boston and his colleagues have used it to identify a possible way to counter drug resistance in acute lymphoblastic leukaemia. The database has also been used to reveal how certain compounds work, through similarities between the gene-expression profiles they evoke and those of compounds of known activity. Both findings are described in papers in *Cancer Cell*.

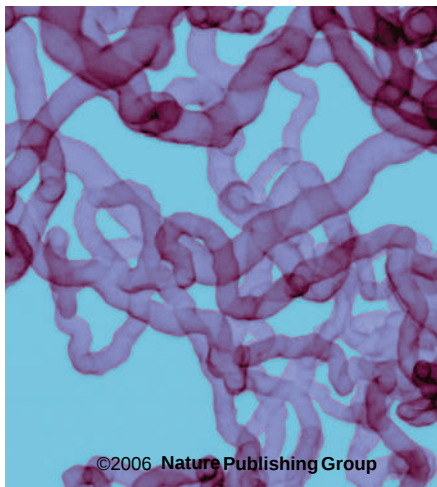
The team proposes expanding the database to feature signatures from cells treated with all approved drugs, and those from cells in which certain genes have been switched off.

NANOTECHNOLOGY**Sort it out**

Nature Nanotech. **1**, 60–65 (2006)

Carbon nanotubes (pictured left) can be sorted by size and electronic type thanks to a method developed by chemists in the United States. Tubes of varying properties tend to be synthesized tangled up together.

The technology, developed by Mark Hersam's group at Northwestern University in Evanston, Illinois, is based on the centrifuge, in which rapid spinning causes objects of different density to separate. Its creators believe commercial opportunities abound for the method, because it can be easily scaled up.



PHOTOTAKE INC./ALAMY

The researchers added surfactants to separate the tubes by electronic type — sorting those with metallic properties from those that behave like semiconductors. The team suggests this works because a tube's conductivity determines how strongly the surfactants cling to its surface, which, in turn, affects density. The process can be repeated to obtain purer samples.

ANIMAL BEHAVIOUR

Thinking ahead

Nature Neurosci. doi:10.1038/nn1781 (2006)

Insight into how barn owls (pictured right) track their prey is provided by Eric Knudsen and his colleagues at Stanford University Medical School, California.

The researchers studied how an owl perceives a moving sound, presented through earphones. A part of the owl's brain known as the optic tectum is arranged to form a map of auditory space, with different neurons responding to sounds coming from different places. Electrical recordings of the activity of these neurons showed that the stimulus elicited a response in neurons corresponding not to the sound's current location, but to where its source would be about 100 milliseconds later.

This type of 'predictive perception' has previously been reported in humans. In owls, it should allow a bird to turn its head to spot the source of a noise — which might become its next meal.

MICROBIOLOGY

A toxin's accomplice

Proc. Natl Acad. Sci. USA doi:10.1073/pnas.0604865103 (2006)

Plants engineered to produce the *Bacillus thuringiensis* toxin kill the insects that feed on them. New research suggests that the toxin works by a previously unsuspected mechanism.

Jo Handelsman of the University of Wisconsin, Madison, and her colleagues found that moths fed antibiotics to sterilize their midguts are immune to *B. thuringiensis*, implying that the insects' natural gut bacteria underpin its potency. Restoring the gut bacteria reversed the effect.

It had been thought that the *B. thuringiensis* toxin might cause starvation, through damage done to cells lining the insect's gut. Instead, the team speculate that this damage triggers sepsis, because gut bacteria leak out. The researchers question whether similar synergies exist in some human diseases.



A. KNUDSEN

CHEMISTRY

Scent packing

J. Am. Chem. Soc. **128**, 11772–11773 (2006)

Isonitriles smell almost supernaturally bad, and one scientist has described their odour as "extremely distressing". As a result, their potential as tools in chemical synthesis has not been fulfilled.

Michael Pirrung and Subir Ghorai at the University of California, Riverside, have now synthesized isonitriles with pleasantly fragrant ester groups attached. These isonitriles retain their useful properties, but lose their stench. The resulting products, say the lab's 'smelling team', have various mild smells, from old wood to taffy.

In addition, the perfumed molecules deliver higher yields than the stinky versions in a common reaction — the Ugi conversion.

PARTICLE PHYSICS

Predictable flips

Preprint at <http://arxiv.org/abs/hep-ex/0609040> (2006)

High-energy physicists have measured the oscillations of a particle that switches rapidly between matter and antimatter, bringing a 20-year effort to an end.

A team at Fermilab's Tevatron particle accelerator in Batavia, Illinois, studied the quick-switch nature of the B_s meson (which is composed of a bottom quark and a strange anti-quark) using a detector known as CDF II to track the decay of the mesons. They calculated that the particle oscillates between matter and antimatter three trillion times per second.

Previous work had only managed to set limits on the oscillation rate. This measurement pins down the properties of the particle, and agrees with the predictions of the standard model of particle physics.

JOURNAL CLUB

Michael Wagner
University of Vienna, Austria

A microbial ecologist discovers another benefit of sex.

Before studying microbiology I mostly associated bacteria with their negative effects, such as disease. So I was fascinated to learn that many bacteria live in eukaryotes such as humans, increasing the fitness or even securing the survival of their hosts. My interest in such symbioses has since shaped research in my lab.

My group investigates, among other things, the interactions between amoebae and their bacterial symbionts. But the research that most recently caught my eye was carried out in aphids — one of the best-studied model systems for symbiotic interactions between animals and bacteria.

Aphids not only harbour primary bacterial symbionts, which produce essential nutrients, but often also play host to secondary symbionts, which confer fitness benefits.

Both types of symbiont are known to be passed to offspring by the mother. But this alone cannot explain infection patterns in aphid populations, which suggest an additional mechanism for the transmission of secondary symbionts.

In a recent paper, Nancy Moran and Helen Dunbar at the University of Arizona in Tucson solved the riddle. They observed that secondary symbionts are abundant in the male aphid reproductive system and can stably infect the female during sex (N. A. Moran & H. E. Dunbar *Proc. Natl Acad. Sci. USA* **103**, 12803–12806; 2006).

Paternal symbiont transfer might be widespread — symbiont cells were recently observed in the sperm of a marine sponge (K. M. Usher *et al. Mar. Freshwater Res.* **56**, 125–131; 2005) and a nematode (G. R. Noel & N. Atibalentja *Int. J. Syst. Evol. Microbiol.* **56**, 1697–1702; 2006). Thus, at least for some animals, it seems that sex is not only about reproduction; it is also an opportunity to gain beneficial microorganisms.

SPECIAL REPORT

Climate in court

A forthcoming case in the Supreme Court could push the United States towards regulating against global warming, says **Emma Marris**.

Is carbon dioxide a pollutant? And, if so, should a reluctant Environmental Protection Agency (EPA) be forced into regulating the gas? These questions will be put before the Supreme Court's nine justices this winter, as they consider the case of the Commonwealth of Massachusetts *et al.* v. EPA.

The case is important for a number of reasons. It is the first of the many lawsuits on climate change to reach the country's highest court. If the EPA loses, it would be forced to contemplate regulations to cap the amount of greenhouse gases emitted by cars. And, most importantly, it may encourage Congress to pass laws regulating carbon dioxide and other gases, lest the courts do its work for it.

It is also a sign of the frustration felt by many US states with the Bush administration's inaction on climate change. States and even cities across the country have entered into their own regional climate agreements and tried to regulate greenhouse gases. California has even signed agreements with the United Kingdom, and on 27 September passed a bill that will return the state's emission levels to those of 1990 by 2020.

The case is being brought by the state of Massachusetts, along with climate warrior California, ten other states, American Samoa and a long list of cities and environmental groups. It stems from a 1999 formal request to the agency that it regulate greenhouse-gas emissions by cars and trucks under the landmark Clean Air Act. The issue has been working its way through the EPA's regulatory system and the courts ever since.

The case is not yet on the Supreme Court's

schedule, but James Milkey, the counsel of record for Massachusetts *et al.*, says it is likely to be heard in "the last week in November or the first week of December". Milkey's side filed its arguments on 31 August; the EPA has applied for an extension to its 5 October filing deadline. But the agency will probably use arguments similar to those it filed when it tried to dissuade the Supreme Court from

"Carbon dioxide is clearly a pollutant — it is the dose that makes the poison."

taking the case (see 'Legal arguments'). Following a day-long hearing, the court will issue an opinion by the end of its term, in late June or early July next year.

Among its arguments, the EPA will probably say that the science on global warming is too uncertain to be the basis of expensive regulations. This threadbare argument has elicited two 'amicus curiae' briefs — documents filed with the court by someone not party to the case. One was from a group of scientists and one from four former EPA administrators. Both say there is a strong scientific consensus, and point out that nothing in science is ever completely certain.

The scientists' brief was filed by 18 researchers, including James Hansen, head of the NASA Goddard Institute for Space Studies in New York, and Nobel prizewinner Sherwood Rowland of the University of California, Irvine. They accuse the EPA of misrepresenting their findings, adding: "The evidence of these changes...has crystallized a remarkable consensus within the scientific community: climate warming is happening, and human activities are very likely a significant causal factor."

Another EPA argument, that greenhouse gases are not pollutants, has met with derision from environmentalists. "Carbon dioxide is



clearly a pollutant — it is the dose that makes the poison," says Andrew Aulisi of the World Resources Institute, an environmental think-tank in Washington DC. "These are political appointees at the top of the agency taking their cues from the White House."

The Supreme Court took the case, experts believe, because of the larger questions it raises over whether a US agency can decide not to do what it is called upon to do by law because of external considerations. In this case, the EPA

D. SAILORS/CORBIS

Action on greenhouse gases

The courts are a popular avenue for those wanting action on greenhouse-gas emissions. If challenges are won, they can force government agencies to do something concrete; even if lost, they publicize the issue. Either way, they light a fire under Congress, which would probably rather write regulations on greenhouse gases

itself. Here are a few key cases: **California v. 6 Automakers.** Filed on 20 September, this complaint sees the greenest US state claiming that car makers' carbon-emitting vehicles are a "public nuisance". A similar case in 2004 against power companies was dismissed, but will be appealed.

Friends of the Earth v. Watson. Environmental groups are suing government agencies for spending money on overseas energy projects without taking into account the environmental impacts of the carbon that will ultimately be emitted. The parties await a decision from the California federal district court.

Center for Biological Diversity v. Kempthorne. The idea is that, under the Endangered Species Act, the government must limit threats from emissions to species such as polar bears and corals. The case was settled this summer when the government agreed to look at the status of polar bears. Green groups hope to use the strategy again. **E.M.**



Twelve US states are suing the Environmental Protection Agency for not regulating greenhouse gases emitted by cars.

has argued that regulating greenhouse gases would interfere with the Bush administration's foreign policy: it would remove the bargaining chip of reducing emissions in exchange for reductions by developing countries.

Madeleine Albright, former secretary of state, criticizes this argument. She has filed an amicus brief saying that the president's foreign policy should not influence agencies directed by Congress. Power thus leaks from Congress to the president — a trend that political observers have kept a wary eye on under George Bush. Albright adds that, in any case, negotiation over emissions reductions has not been the Bush administration's strategy.

"The United States is not actually negotiating with developing countries," says Jody Freeman, who directs the environmental law programme at Harvard Law School in Cambridge, Massachusetts, and drafted the Albright brief. "It is pursuing a volunteerism strategy."

"You strip away the global warming and the question is really 'can they do that?'" she says. For this reason, many believe Massachusetts could be favoured by the current judges, despite

Legal arguments

Standing

For the judges to accept the arguments of Massachusetts *et al.*, the group must prove that it has standing — in other words, that it is the injured party, and that a decision in its favour will redress the injury. Massachusetts says it is already seeing the effects of climate change — for example, the resulting sea-level rise threatens the coastal state. The Environmental Protection Agency (EPA) is likely to say that the injury is not caused directly by its decision not to regulate greenhouse gases, that regulation would be unlikely to significantly roll back climate change, and thus that the injury cannot be fixed by a decision in Massachusetts' favour.

Does the EPA have authority to regulate greenhouse gases under the Clean Air Act?

The EPA says not, reasoning that if Congress wanted it to regulate greenhouse gases

— a politically contentious and economically significant action — it would have said so explicitly. The agency notes that when Congress wanted it to regulate pollutants that thinned the ozone layer, language was added to the Clean Air Act to say so. Massachusetts *et al.* say arguments about Congress's intent are not legally robust — the agency must conform to the law as it is written. The act defines a pollutant as "any air pollution agent... including any physical, chemical, biological, radioactive substance or matter which is emitted into or otherwise enters into the ambient air". Massachusetts *et al.* say this describes greenhouse gases perfectly. The EPA notes that, strictly speaking, the definition requires a pollutant also to be an "air pollution agent", and it has come up with various reasons why

greenhouse gases should not count as such.

If the EPA does have the authority to regulate greenhouse gases under the act, can it decide not to do so?

Massachusetts *et al.* think not. The law says the agency "shall" regulate pollutants "which may reasonably be anticipated to endanger public health or welfare". But the EPA says it can choose even whether to determine if a pollutant is a threat. The EPA hasn't judged greenhouse gases on this and says it won't do so because the science is still too uncertain. It says regulation would undercut US foreign-policy manoeuvres that might bargain over greenhouse gases with other nations, and that regulation would amount to setting miles per gallon limits, which is already the job of the transportation department. **E.M.**

their conservative leanings. "Even justices who may not be particularly sensitive to the need to regulate global warming may think that an agency is overreaching," says Freeman.

If Massachusetts wins, the EPA will not immediately regulate car emissions. Agency scientists would merely have to determine formally whether greenhouse gases are likely to endanger public health. If they decide this is the case, the agency may begin hammering out regulations. But all this would take years to unfold.

More to the point, the case may push Congress to pass some climate-change laws. "Part of all this litigation that's going on in various places is to try to force the issue onto the congressional agenda," says Deborah Sivas, director of the Environmental Law Clinic at Stanford University in California, and drafter of the brief by the four former EPA administrators (see 'Action on greenhouse gases').

While it waits for the court case to play out, the EPA has remained mute on California's request that it be allowed to regulate greenhouse-gas emissions from cars at state level

"Even justices who may not be particularly sensitive to the need to regulate global warming may think that an agency is overreaching."

— cutting them 30% by 2016. Thanks to a historical quirk in the Clean Air Act, California is the only state allowed to set different standards from those set by the federal government — with the agency's permission. But once California changes its standards, any state can follow. A letter released on 26 September, signed by more than 100 members of Congress, called on EPA administrator Stephen Johnson to stop stalling. "There is no basis for EPA to treat this request differently" from the many other requests it has granted, the letter said.

Almost a quarter of the members of Congress signed that letter. So perhaps the rumour in green circles that a bill addressing climate change could pass in the next session of Congress is well founded. Much depends on the November elections, in which all congressmen and a third of senators will fight their seats. If the Democrats gain a majority in the House of Representatives — a possibility at least — they would control the committee where climate-change bills will come up, and things may really start to move.

For more on the impact of the mid-term elections, see next week's issue of Nature.

CLIMATE CHANGE IN FOCUS

From melting ice to carbon markets, find all our stories on global warming.

www.nature.com/news



GETTY

ON THE RECORD

“That’s one small step for a man, one giant leap for mankind.”

Computer analysis shows that Neil Armstrong, the first man on the Moon, didn’t fluff his lines after all.

“If you had told me I would have death threats in a case like this, I would have said you were crazy. But I did.”

Judge John Jones is stunned by what happened when he struck down the decision of a school board in Dover, Pennsylvania, to teach intelligent design in schools.

OVERHYPED

Allergy-free kittens

Allerca of San Diego, purveyor of “the world’s first scientifically proven hypoallergenic cats”, is making much of the fact that its kittens were bred naturally rather than using transgenic methods, all in a mere 18 months. How did the firm avoid years of selective breeding? It stumbled across three cats with mutated allergen

proteins during pre-testing for the gene-silencing method it was preparing to use, neatly sidestepping most of the hard work. And how did

Allerca obtain “independent confirmation” that the \$4,000 cats are for real? Ten blindfolded volunteers, stroking experimental or control cats. The research has yet to find a publisher.

SCORECARD



Endurance
Mars rover Opportunity caused joy and panic as it reached 1,000 days of exploring — it was built to last just 90 days, so the onboard software could only accommodate dates with 3 digits.



Astronomy
Iceland’s capital Reykjavik was plunged into darkness last week as lights were switched off to give locals an unprecedented view of the night sky, despite a few clouds.

Sources: *Kansas.com*, *Reuters*

Youthful duo snags a swift Nobel for RNA control of genes

The call from Stockholm was received by two unusually young scientists, an impressively short time after they demonstrated a fundamental control of gene expression.

The 2006 Nobel Prize in Physiology or Medicine has gone to Andrew Fire, now at Stanford University, California, and Craig Mello, now at the University of Massachusetts Medical Center in Worcester. Both are in their forties. It honours a discovery that has transformed biological research and may, in the future, prove useful in treating human disease.

The discovery is called RNA interference, or RNAi. When a gene is to be expressed, it sends instructions to the cell’s protein synthesis machinery. The intermediary is messenger RNA (mRNA), which has a structure complementary to that of the gene. In their breakthrough paper, published in *Nature* in 1998, Fire and Mello, and colleagues, demonstrated that these mRNAs can be targeted for destruction by specific double-stranded forms of RNA (A. Fire *et al.* *Nature* 391, 806–811; 1998).

Researchers already knew that ‘antisense’ RNA — an artificial molecule whose sequence complements the mRNA — could silence specific genes when taken up by a cell. But the effect was modest and inconsistent. And, confusingly, the same effect was seen with ‘sense’ RNA.

In a series of simple and elegant experiments on a muscle gene in the nematode worm *Caenorhabditis elegans*, Fire and Mello showed that a powerful and consistent effect required the sense and antisense RNAs to be stuck together, as double-stranded RNA. When injected with the double-stranded RNA, the worms twitched awkwardly, just like mutant worms lacking the muscle gene. The researchers also showed that mRNA was destroyed by the treatment, rather than being masked as others had believed. And they showed that the double-stranded RNA can cause more copies of itself to be made, can spread between cells and can even be inherited by progeny.

In the *Nature* paper the researchers speculated that organisms might use this mecha-



Silence is golden: Craig Mello (left) and Andrew Fire.

nism to control expression of their genes — a fact soon shown to be true. RNAi turned out to be an important way of controlling ‘jumping genes’, which can insert themselves throughout the genome and disrupt gene function. It is also thought to help protect against viral

infections, at least in simple organisms.

Fire and Mello’s subsequent research, and that of the many researchers who flooded the field after them, have since filled in the details. The RNAi pathway has been fully worked out and shown to apply to all genes — not just in worms but in nearly all other species.

“There is hardly a lab not using the tools of RNAi to turn off the genes that they are studying,” says Tom Tuschl of Rockefeller University, New York, who joined the RNAi elite soon after the Fire and Mello discovery. RNAi reagents have been used to find genes involved in everything from ageing to cancer. Clinical hopes are pinned on finding ways to use RNAi to reduce the activity of genes that cause disease.

Most Nobel prizes are given many years after the relevant discovery. The Fire and Mello award, given just eight years after publication of their paper, is reminiscent of Kary Mullis’s 1993 chemistry Nobel. That prize was awarded for his 1985 invention of the polymerase chain reaction — a method of gene amplification that invaded research labs just as fast and comprehensively as the RNAi technique has.

Ronald Plasterk of the University of Utrecht in the Netherlands, who worked with Fire when they were young researchers at the Laboratory of Molecular Biology in Cambridge, UK, says he felt that a Nobel was around the corner. But he didn’t think it would be so soon. For a laugh, on the morning of the announcement Plasterk asked his postdoc to call Fire in a pronounced Swedish accent. “But then we heard he had really won, so the joke had to be dropped.”

Mello seemed stunned when interviewed on the Nobel webcast, not long after he had heard the news. “I seem too young,” he said. “And isn’t the gap unusually short?”

Alison Abbott

STANFORD UNIV.; UNIV. OF MASSACHUSETTS

Cosmic ripples net physics prize

The discovery that finally turned cosmology into a hard science has scooped this year's Nobel Prize in Physics. John Mather and George Smoot's measurements of the earliest moments of the Universe, made with the COBE (Cosmic Background Explorer) satellite, created a whole field of research and captured the imagination of millions around the world.

"We used to have derision cosmology," says Max Tegmark, a cosmologist at the Massachusetts Institute of Technology. "People used to make fun of this field as being very flaky and philosophical. Now we have precision cosmology."

COBE's goal was to study the heat left from the Big Bang. The existence of this afterglow, known as the cosmic microwave background, was predicted in the 1940s and confirmed by measurements made in the 1960s.

The Nobel prize is awarded for two discoveries. The first, from a team led by Mather, now at NASA's Goddard Space Flight Center in Greenbelt, Maryland, is the measurement of the microwave background's spectrum, which confirmed that it had the shape of 'black-body' radiation. Such radiation is emitted by an object of a particular temperature — in this case, by the Universe when its temperature was thousands of degrees hotter than today.

This result was the first indication that cosmologists would be able to use properties of the radiation to test theories about the origin of the Universe. Mather showcased the findings at a conference in January 1990 and was greeted with a standing ovation. "In my whole career, it's the only one I've witnessed," says Carlos Frenk, a cosmologist at the University of Durham, UK, who has worked with COBE data.

Smoot, now at Lawrence Berkeley National Laboratory

in California, was in charge of seeking slight fluctuations in the temperature of the microwave background. These are the imprint of the Universe's early structure — the clumps of matter that formed the seeds of today's galaxies and stars. Detecting this anisotropy was a feat that some thought beyond the sensitivity of COBE. "A lot of people thought they weren't going to find anything," says Tegmark.

But in 1992 Smoot announced to worldwide acclaim that he had found the ripples. This was the strongest evidence so far that the Big Bang happened. "Anisotropy is like looking at the fossil record and having that exposed for the first time," says Anthony Lasenby of the University of Cambridge, UK.

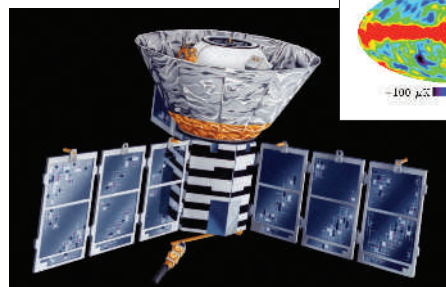
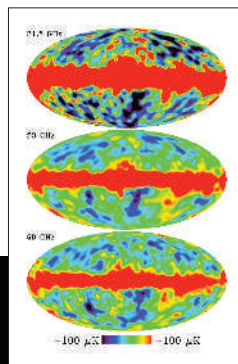
At the time, Lasenby was very close to finding similar results of his own. "We were working on an experiment in Tenerife," he says. "It was showing the first signs of fluctuations but we needed more time." Smoot got there first. But it was important to verify and correlate Smoot's results, says Lasenby, without a hint of bitterness.

After COBE came WMAP (the Wilkinson Microwave Anisotropy Probe), named after David Wilkinson, who was instrumental in setting up the COBE mission and who died in 2002. Many believe he would have shared today's prize had he survived. WMAP surveyed the background radiation in even greater detail, and will be followed in 2008 by the European Space Agency's Planck satellite.

But the impact of COBE is so far unmatched. Physicist Stephen Hawking caught the mood in 1992 when he described Smoot's results as "the greatest discovery of the century — if not of all time". There was controversy too, when Smoot said that looking at the observations was like looking into the face of God. "Bringing God into science is not always a good thing," says Frenk.

Mather and Smoot are known to have clashed personally, says astronomer George Efstathiou at the University of Cambridge, and fell out over the way the anisotropy results were released. "But maybe now that their names will be forever linked together as co-prizewinners, they might resolve their differences."

Katharine Sanderson and Jenny Hogan
For details of the 2006 Nobel Prize in Chemistry, see www.nature.com/news



Glowing report: the COBE satellite produced data (inset) that shed light on the early Universe.



JAVA MUD VOLCANO SEEMS UNSTOPPABLE
Could Indonesia's mud flow be put to good use?
www.nature.com/news

CRISP

Theorists snap over string pieces

Two recently published books are riling the small but influential community of string theorists, by arguing that the field is wandering dangerously far from the mainstream.

The books' titles say it all: *Not Even Wrong*, a phrase that physicist Wolfgang Pauli used to describe incomplete ideas, and *The Trouble with Physics: The Rise of String Theory, the Fall of a Science, and What Comes Next*. Both articulate a fear that the field is becoming too abstract and is focusing on aesthetics rather than reality. Some physicists even warn that the theory's dominance could pose a threat to the scientific method itself (see page 507).

Those accusations are vehemently denied by string theorists, and the books — written by outsiders — have stirred deep resentment in the tight-knit community. *Not Even Wrong* was published in June and *The Trouble with Physics* came out in September; shortly after they appeared on the Amazon books website, string theorist Luboš Motl of Harvard University posted reviews furiously entitled “Bitter emotions and obsolete understanding of high-energy physics” and “Another postmodern diatribe against modern physics and scientific method”. As *Nature* went to press, the reviews had been removed.

Few in the community are, at least publicly, as vitriolic as Motl. But many are angry and struggling to deal with the criticism. “Most of my friends are quietly upset,” says Leonard Susskind, a string theorist at Stanford University in California.

String theory postulates that the Universe consists of tiny strings vibrating in ten or so dimensions. Its fortunes have been buoyed by popular books in the past — most notably

Brian Greene's 1999 bestseller *The Elegant Universe*, which said that the approach might unify the incompatible theories of gravity and quantum mechanics.

Strung up

But the theory has its share of problems, and these are the focus of the new works. For one thing, recent calculations suggest that it generates 10^{500} possible models of the Universe (see *Nature* 439, 10–12; 2006). This renders the theory essentially meaningless, according to critics. When these countless possibilities were first announced, Lee Smolin was already working on the book that eventually became *The Trouble With Physics*. A physicist at the Perimeter Institute for Theoretical Physics in Waterloo, Canada, Smolin had previously worked on string theory. “I and many other people began to get worried,” he says.

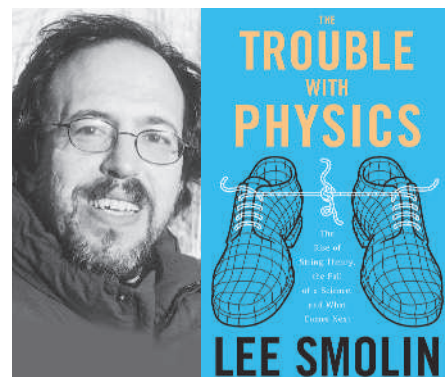
Another difficulty is that if strings exist, they would be detectable only at energies far above anything that today's experiments can measure. “As you're not constrained by having to reproduce experiments, you can go off and play with whatever you want,” says Peter Woit, a mathematician at Columbia University in New York City, and author of *Not Even Wrong*. “The danger is that you'll end up with the theoretical community becoming completely isolated from the rest of physics.”

Smolin, whose book promotes an alternative theory known as loop quantum gravity, adds that string theorists have intentionally cut themselves off. “None of the major string theory groups has hired a postdoc or faculty member working in any of the other approaches to quantum gravity,” he says. “But other research groups in quantum gravity have often hired young people working in string theory out of a sense that it should be encouraged.”

Boundary issues

String theorists dispute the claim that they are isolating themselves. In recent years the theory has contributed significantly to heavy-ion physics, according to Joe Polchinski, a string theorist at the Kavli Institute for Theoretical Physics in Santa Barbara, California. When the Relativistic Heavy Ion Collider at Brookhaven National Laboratory in Upton, New York, first produced a hot quark gas, it was string theory that correctly predicted, retrospectively, some of the gas's properties. “In many ways, I feel the boundaries with other areas of physics are coming down,” Polchinski says.

Warren Perkins, a cosmologist and particle



D.F. GRASER

“None of the major string theory groups has hired someone working in any of the other approaches to quantum gravity.”

theorist at Swansea University in Wales, agrees. In recent years, he says, string theory has been proven equivalent to a conventional ‘field theory’, of the type being used to predict how particles will behave at the soon-to-open Large Hadron Collider, sited near Geneva. “There has been a lot of cross-talk between field theory and string theory,” says Perkins. In some cases, string theory has provided far simpler approximations than its field-theory counterpart.

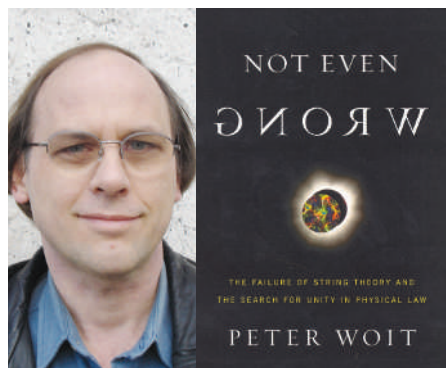
But strings have yet to provide the elusive link between gravity and quantum mechanics, hoped for by so many theorists. “The claims, when it comes to theoretical physics, tend to be exaggerated,” says Abhay Ashtekar, who works on quantum gravity at Pennsylvania State University in University Park. He believes the inability of the community to live up to those expectations has made it defensive.

The books leave string theorists such as Susskind wondering how to approach such strong public criticism. “I don't know if the right thing is to worry about the public image or keep quiet,” he says. He fears the argument may “fuel the discrediting of scientific expertise”.

That's something that Smolin and Woit insist they don't want. Woit says his problem isn't with the theory itself, just some of its more grandiose claims. “There are some real things you can do with string theory,” he says.

Smolin agrees, and says he hopes theorists will read his book to get a better understanding of his specific issues. “If they don't want to buy it, tell them to get in touch with me and I'll send them a copy.”

Geoff Brumfiel



P. CRUZ

“The danger is that you'll end up with the theoretical community becoming completely isolated from the rest of physics.”

Q&A: Fabio Mussi

VENICE

Before last April's general election in Italy, *Nature* reported deep dissatisfaction among scientists about what they considered disastrous science-policy decisions by Silvio Berlusconi's right-wing government. An international conference on evolution in Venice last month was opened by new research minister Fabio Mussi — a member of the Democrats of the Left party, part of the ruling coalition. Known for his bluntness and quick humour, he gave **Alison Abbott** a snapshot of his plans for putting things right.

Q What is the most important issue?

A Money money money. Italy spends less on science and innovation than it should. We want to increase public investment. We want Italian scientists to make better use of opportunities to win European Union research funds. And we want to find ways to make industry contribute more.

Q What can be done about the much-criticized leadership of organizations such as the Italian Space Agency and the National Research Council?

A It's another very important priority, to change a few presidents of our research organizations. It hasn't proved very easy, but when we have done

this, other problems will be much easier to solve.

Q What about the Italian Institute of Technology, conceived by the last government as a version of the Massachusetts Institute of Technology and viewed by many as a white elephant?

A This was not born well, but I cannot struggle with my conscience to kill it. It was given €240 million (US\$305 million) by the last government that it hasn't spent — we are not going to give it any more money until it does so. I am very much in favour of merging it with the European Institute of Technology [planned by the European Union along similar lines].

"It seems to me that those who believe in intelligent design are the ones who support politics that are not intelligently designed."



New wave: research minister Fabio Mussi opened an international conference on evolution in Venice last month.

Cloners break away from the herd



B. CURTIS/PA/EMPICS

Following on from Dolly: using differentiated cells might be key to efficient cloning.

Surely it's easier to clone mammals from cells that are dividing, than from those fated never to do so again? That's certainly the assumption with which most biologists have been working.

But results from a study in mice suggest that the opposite may be true: in attempts to clone mouse embryos from three groups of blood cells, the most differentiated ones worked best. The finding could improve the lamentable state of cloning, although experts are stumped for an explanation.

In cloning, researchers insert the nucleus of a cell, say a skin cell, into an egg stripped of its own nucleus. This process is thought to reprogramme the DNA in the nucleus, giving it the potential to orchestrate the development of an embryo. Nuclei from stem cells, which can give rise to various other types of cells, were assumed to be easier to reprogramme than nuclei from differentiated cells that have lost the ability to divide. Cloning from embryonic stem cells is much more successful than from adult stem cells, for example.

Now Jerry Yang of the University of Connecticut, Storrs, and his colleagues have compared mouse blood cells at three stages of differentiation. One group comprised stem cells that could produce all blood-cell types. A second group of cells could make only a hand-

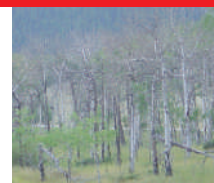
ful of cell types, and a third group was composed of completely differentiated white blood cells called granulocytes.

The granulocytes were easiest to clone. With these cells, 35% of cloning attempts yielded young embryos, compared with 11% in the intermediate group and less than 8% from the stem cells (L.-Y. Sung *et al. Nature Genet.* doi:10.1038/ng1895; 2006).

Scientists have long debated whether cloning is so difficult because reprogramming doesn't work well, or because the few clones produced grew from rare stem cells lurking among the differentiated cells used. Dolly the sheep, hailed as the first mammal cloned from an adult cell, was created from a mammary gland — but some claim she could have come from a stem cell in that tissue.

Previous attempts to clone from fully differentiated cells have been problematic, and it's possible that granulocytes may simply contain particular biochemical factors that make them easy to clone. But if Yang's result holds generally, Dolly may have come from a differentiated cell after all. It would also be good news for the technique of therapeutic cloning, in which researchers hope to clone patients' cells and use the resulting embryos to extract embryonic stem cells to treat disease. ■

Helen Pearson



WHERE HAVE ALL THE ASPEN GONE?

The mysterious deaths of trees in the western United States.

www.nature.com/news

Wikipedia rival calls in the experts

The current incarnation of Wikipedia is both phenomenally successful and, in the eyes of some critics, fundamentally flawed. The online encyclopaedia now includes more than a million entries in English alone. Although anyone can edit any article, its accuracy, at least on science topics, is surprisingly high. But Wikipedia has never given experts special standing when it comes to determining content. And that, critics say, deters the people who ought to be contributing from doing so.

Just how big a drawback that is will now be tested, with the launch of an online encyclopaedia that will give privileged status to scientists and other experts. Citizendium, a pilot version of which is due to go live in the next week, will use

all of Wikipedia's content but will host it at another website (<http://citizendium.org>) and edit it differently. Editors with appropriate academic qualifications will have the power to settle disputes about wording, for example, and stamp articles they perceive to be accurate as 'approved'.

"One reason we are setting this up is to give scientists and other scholars a new organizational framework to clean up and improve on the work started by Wikipedia," says Larry Sanger, a philosopher and co-founder of Wikipedia, who is the driving force behind Citizendium. "Wikipedia is

now the first stop for many people in their search for information on scientific topics. Many scientists would like to help make sure this resource remains accurate, but they have no desire to navigate the treacherous waters of Wikipedia's editorial system, which accords them no official role."

Reactions from the many bloggers who track the progress

"Many scientists have no desire to navigate the treacherous waters of Wikipedia's editorial system."

of Wikipedia have been mixed. Some say Sanger will struggle to define what constitutes expertise, and that arguments about content will be replaced by arguments over who is or isn't an expert. Others think Citizendium

will find it hard to attract regular contributors, as the prospect of having an edit overruled by a higher power will not appeal.

But scientists who contribute regularly to Wikipedia say Citizendium has promise. "I like the idea notionally," says Vaughan Bell of the Institute of Psychiatry in London, who contributes to Wikipedia's schizophrenia page, among others. William Connolley of the British Antarctic Survey in Cambridge, UK, who updates Wikipedia climate entries, adds that some scientists have become frustrated with Wikipedia because of the difficulty in agreeing edits, although both he and Bell agree that conflict can sometimes result in better articles.

Jim Giles

Scientists push for stronger voice in Congress

Hoping to influence the outcome of the 7 November US midterms, a group of scientists is trying to sway local elections over science issues.

The group, Scientists and Engineers for America, is advancing a 'scientific bill of rights', which seeks to protect government scientists from political interference. Already the group boasts three Nobel laureates: biological chemist Peter Agre, physicist Burton Richter and biologist Alfred Gilman. Its board also includes Neal Lane and John Gibbons, both former White House science advisers under President Bill Clinton.

Unlike other issue-based organizations, the group will also directly endorse congressional candidates, depending on their view on topics such as stem cells and climate change. "If there is an elected member who denies that global warming exists," says board member Michael Stebbins, "then it would be our belief that that person should no longer have their job."

Analysis shows no shortcomings for hobbit

There's a riposte in the debate over the origins of the tiny 'hobbit' hominids thought to have lived on the isolated Indonesian island of Flores until about 12,000 years ago. Some groups have disputed the original description of the hobbit as the new species *Homo floresiensis*, claiming instead that the remains might be a dwarfed *Homo erectus*, a pygmy, a deformed modern human or even a holdover from a much more ancient hominid lineage.

Debbie Argue of the Australian National University in Canberra and her colleagues have now tested these assertions in a series of analyses — and state that the hobbit is, indeed, a distinct species.

Their results (D. Argue *et al.* *J. Hum. Evol.* 51, 360–374; 2006) largely reiterate the findings in the original paper describing the discovery of the hobbit — although

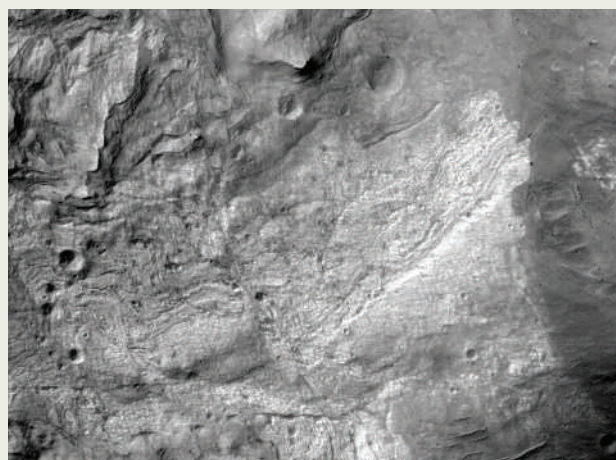


Findings suggest that the tiny hobbit skull (left) is not the deformed head of a human.

Cracking close-up of Mars

The latest spacecraft orbiting Mars can now spot objects as small as 75 centimetres across. This image is the first close-up photo sent back from the most powerful telescopic camera ever sent to another planet — the HiRISE instrument aboard the Mars Reconnaissance Orbiter.

The image, taken on 29 September, shows layers of rock, possibly laid down through sedimentary processes, in the Ius Chasma region of the Valles Marineris system of canyons. HiRISE scientists at the University of Arizona, Tucson, will start full-time science studies in November, focusing on photographing the north polar region of Mars where the Phoenix mission is set to touch down in May 2008.



NASA/JPL/UA

they go into more detail. They affirm that *H. floresiensis* was a distinct species, an offshoot of an early migration of archaic hominids from Africa, that followed its own evolutionary path.

The analyses rule out any association with normal *H. sapiens*, including pygmies, and speak strongly against freak abnormalities such as unusually small heads.

SKA telescope placement down to two-nation race

China and Argentina are out of the running, leaving South Africa and Australia as the final contenders to host the enormous radio telescope called the Square Kilometer Array (SKA).

Last week, the international committee coordinating the project reduced the number of possible locations for SKA. South Africa would place SKA's centre in the Karoo desert; Australia would build it in the west. The host country is likely to be decided towards the end of this decade. When completed, SKA will consist of thousands of small dishes and antennas arranged in clusters over an area 3,000 kilometres across.

Construction is slated to begin in 2011, with operations to start in 2014.

Representatives call for NIH structural overhaul

The House of Representatives has taken steps towards an overhaul of the US National Institutes of Health (NIH). Last week, the House overwhelmingly passed a 'reauthorization' bill laying out structural and administrative changes for the NIH.

The bill includes, for instance, the establishment of a 'common fund' for research that bridges NIH institutes. The money would come from a variety of sources and would be capped at a maximum of 5% of the NIH budget. But for the bill to become law, it would have to be approved by the Senate late this year — an uncertain prospect at best. Lawmakers are also not compelled to abide by the bill's request for budget increases of up to 5% for the coming three years.

"It's an important statement," says Howard Garrison, policy director at the Federation of American Societies for Experimental Biology in Washington DC. "But it's not money in the bank."

Separately last week, the House also passed a bill that would create an agency to oversee biodefence research. That bill, too, would need to be passed by the Senate to become law.

Plasma milestone at Chinese tokamak plant

China has successfully obtained the first plasma ever confined inside a fully superconducting tokamak (see *Nature* 442, 853; 2006).

On 26 September, after a month-long delay caused by a glitch in one of the turbines, researchers at the Institute of Plasma Physics of the Chinese Academy of Sciences confined a plasma discharge for 2.8 seconds. Jiangang Li, director of the institute, says the numbers were higher than they had expected, but still far below the final goal of achieving a 1,000-second duration. That would make their tokamak by far the best machine for confining plasma for study.

BUSINESS

Copcats gear up to dog biotech brands

Pressure is mounting on US regulators to create an approval track for generic versions of biotechnology drugs. **Meredith Wadman** reports.

Bruce Downey gets exasperated when he looks at how long people have been dithering about the right way to deal with generic copies of biotechnology drugs.

Last month in Washington DC, this chief executive of Barr Pharmaceuticals, New Jersey, was speaking at a meeting of makers of generic drugs. He produced a slide entitled 'Chronology of Delays', which charted seven years of broken promises in the US Food and Drug Administration (FDA) plan to approve generic versions of biologics, the complex medicines made by the biotechnology industry. "Where is the FDA on this?" Downey demanded to know. "They are missing in action."

But things may at last be looking up for Downey, whose company is set to become the world's third-largest generic drugmaker — with annual revenues of about US\$2.4 billion — when it acquires Pliva Pharmaceuticals of Zagreb. A bill was introduced in the US Congress on 29 September that, if passed, would give the FDA the authority to approve applications to sell generic biological drugs — or biosimilars, as they are sometimes called.

The bill's authors include Senator Hillary Clinton (Democrat, New York) and Representative Henry Waxman (Democrat, California). Waxman wrote the 1984 law that established the FDA's existing approval system for generic drugs, but that law does not apply to biologics.

The new bill is unlikely to get very far before Congress adjourns for the year. But its introduction is a warning shot for the US biotechnology industry, which has been able to operate so far without worrying about competition from generics. That arrangement is coming under pressure as the original patents on biotech drugs begin to expire.

"Under our current law, biologic products are effectively given a near-complete monopoly," Waxman told the meeting of the Generic Pharmaceutical Association on 19 September. "The time has come to break this monopoly. Congress can no longer stand by and watch as our reliance on biologics increases, and the cost of these medicines continues to soar."

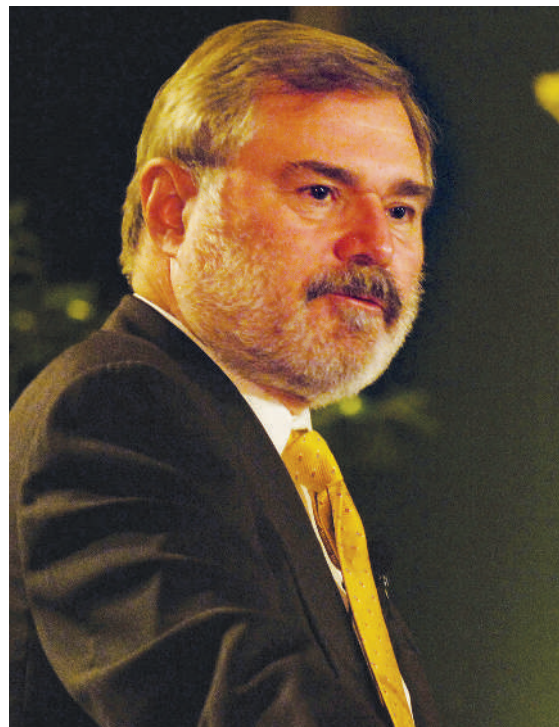
It's not only Democrats who share his outlook. Senator Orrin Hatch (Republican, Utah), who coauthored the original generics law with Waxman, told the meeting that any new law must carefully balance incentives for both brand-name makers and the generics industry. He plans to introduce his own legislation to allow generics in 2007. "Many in the biotechnology industry believe that the creation of a fast-track approval process for off-patent biologics is a terrible idea, and on top of that not doable," Hatch says. "That's where I start getting frustrated."

The industry disputes the technical feasibility of legislation permitting biogenerics. It points to the enormous size, complexity and inherent variability of biologic medicines, which are made in living systems and thus are not exactly reproducible in the way that small-molecule drugs are. Erythropoietin, a hormone that boosts red blood cells, is one example of a popular biological drug. Sold by California-based Amgen as Epogen, it consists of a protein plus a sugar, with a molecular weight of about 34,000. In contrast, the molecular weight of the antidepressant Prozac — fluoxetine hydrochloride — is just 346.

Because of this, brand-name biologics makers say, extraordinary caution is needed in crafting any law intending to speed biosimilars to market. They contend that such efforts are fraught with intellectual-property conundrums, safety concerns and scientific obstacles — key among them is the difficulty of ascertaining whether a generic molecule is close enough to the original medicine to be substituted for it.

"They can write whatever legislation they want, but if the science doesn't allow it, there's an expectation that there will be generic biological products that just isn't going to be met," Amit Sachdev says. Sachdev is an executive vice-president at the Biotechnology Industry Organization (BIO), a Washington-based lobby group. Until last year, he was a senior policy official at the FDA.

Val Giddings, a Washington consultant and former BIO vice-president, says that the generics makers "would like to be able to take shortcuts so they don't have to go through a



Frustrated: Bruce Downey, a maker of generics, is demanding action from US drug agency.

seven-year, \$800-million safety assessment". But that may not be possible. In any case, Giddings doubts that the Waxman bill represents anything but the most preliminary noise in what is likely to be a battle that lasts for years.

There is big money at stake. By the end of this year, biologics with combined annual sales of \$11.5 billion will have come off patent. They include blockbuster drugs such as Epogen, which had sales worth \$2.5 billion in 2005, and Eli Lilly's Humulin, an insulin to treat diabetes that, despite losing patent protection in 2002, logged \$1 billion in sales last year.

Those numbers alone are going to make it difficult for the industry to stall legislative action, at least in the long term, says Gillian Woollett, chief scientist at Engel & Novitt, a Washington law firm that represents biotechnology companies. "When you get into numbers starting with B, people start to pay attention," she says. "It's time for the biotech industry to accept the consequences of its own success."

European push

Movement in Europe may also serve to increase the pressure for change in the United States. The European Union last year opened a regulatory path for the approval of biosimilars (see *Nature* **438**, 154–155; 2005). This year, the European Medicines Agency approved the first two such drugs, generic

"Congress can no longer stand by as the cost of these medicines continues to soar."



J. AUGUSTINO

versions of human growth hormone, and more approvals are on the horizon.

But some economists question the underlying assumption driving acceptance of biogenerics — that access to them will slash prices for consumers. Generics of ordinary drugs commonly become available for one-twentieth the price of the original, or even less. However the costs of actually producing biogenerics will be far higher. And because they will be not quite identical to the originals, regulators are likely to demand at least some clinical trials for biogenerics. In Europe, these are estimated to cost tens of millions of dollars for each drug, against the couple of million commonly required to prove that a generic drug can be substituted for an original medicine.

What's more, because biologics often treat cancer and other grave diseases, it is not clear how willing insurers would be to push patients toward generic versions as a means of cost control. It also remains to be seen whether seriously ill patients would accept substitute products.

Henry Grabowski, an economist at Duke University in Durham, North Carolina, argues that few companies will attempt to make biogenerics and that dramatic price drops will thus be unlikely, at least in the short term (H. Grabowski, I. Cockburn and G. Long *Health Aff.* 25, 1291–1301; 2006).

In the case of ordinary, small-molecule drugs, Grabowski says, "it is important to remember that the rapid pace of generic entry and penetration took many years to evolve. This will also be true for biologics." ■

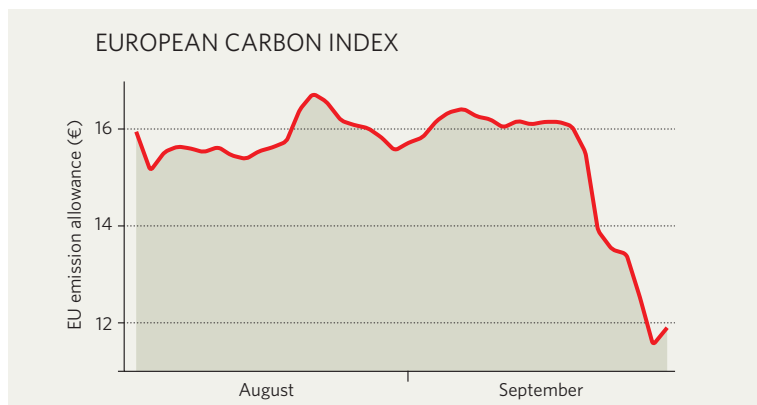
IN BRIEF

SUITOR IDENTIFIED Serono is to be bought by Merck in a deal that values the Swiss biotechnology company at 16.1 billion Swiss francs (US\$12.9 billion). The largest biotech company in Europe, whose best-selling product is a drug for multiple sclerosis, Serono has been scouting for a partner or buyer since last autumn (see *Nature* 438, 557; 2005). Merck of Darmstadt, Germany — unrelated to the New Jersey company of the same name — earlier this year lost out in a bid to buy its rival Schering (see *Nature* 440, 603; 2006). It says Serono will retain its own laboratories and headquarters in Geneva.

MARKED CONFIDENTIAL Icelandic genetics database company DeCode Genetics is suing five of its former staff and their new employer, the Children's Hospital of Philadelphia, for allegedly misappropriating its trade secrets. The suit, filed at the district court in Philadelphia on 26 September, claims that the five sent the hospital proprietary information about DeCode's methods and business plan before leaving to work there. The hospital, which denies wrongdoing, is attempting to build a database of 100,000 children and mine it for information about susceptibility to disease.

IN THE SOUP Christopher Evans, Britain's most prominent biotechnology entrepreneur, has been arrested in connection with a police investigation into loans made to the ruling Labour party. In a statement, Evans said he was "shocked and dismayed" at the arrest. He says he lent the party £1 million (US\$1.9 million) openly, on a commercial basis, and that interest would be paid on it. Critics charge that Labour secretly accepted several large loans just before the 2005 general election in exchange for honours or influence.

MARKET WATCH



SOURCE: EUROPEAN ENERGY EXCHANGE

After steadying in the summer, prices for European Union allowances for emissions of carbon dioxide have dipped precipitously in the past two weeks, to less than €12 (US\$15) for a tonne of emissions.

According to carbon-market analysts, the drop reflects recent falls in oil and gas prices. These encourage electricity generators to switch from coal to natural gas, whose more efficient burning produces less carbon dioxide. The fuel switch has lowered the demand for emission allowances traded at five European carbon exchanges, including the European Energy Exchange in Leipzig, Germany (see chart).

But Nils Medenbach, an emissions-trading specialist at the Frankfurt-based consultancy 3C Climate Change Consulting, says the low prices also indicate that many energy companies think they already have enough

allowances to cover the first period of the carbon-trading scheme — which lasts until the end of next year.

As the year draws to a close, and industries get a better picture of their overall 2006 energy use, large facilities throughout Europe are also tending to sell options rather than buy them, says Milo Sjardin, an analyst at the London-based New Carbon Finance.

Some analysts think that demand for allowances will continue to slip. "You may want to sell now," says Sjardin. "Unless industrial production suddenly increases to very high levels, or drought leads to a dramatic drop in hydropower supply, prices may go down to €5 by the end of the year." Prices are higher, however, for the next trading period, from 2008 to 2012, and market volume remains healthy, dispelling suggestions that the trading scheme is failing altogether. ■

Quirin Schiermeier

Four fires broke out in San Diego on 19 September 2003. Banners were found at each site claiming that the Earth Liberation Front was responsible.



IN THE NAME OF NATURE

What drives environmental activists to fire-bomb laboratories?
Emma Marris investigates a radical fringe of the US green movement.

One day in June 1998, three young environmental activists of a radical bent drove from Eugene, Oregon, to Olympia, Washington. Their route took them through some of the loveliest country in the Pacific Northwest: up Interstate 5, through the Willamette Valley, between dark green forested mountains and misty hillside vineyards. Their van shared the road with logging trucks carrying immense trees hung with lichens and mosses.

According to the Federal Bureau of Investigation (FBI), the three were on their way to take part in fire-bombing two research facilities. Things, however, didn't go according to plan. They picked up one friend and lost another to a shoplifting bust. They lost contact with their second vehicle. But something compelled them to set fire to their target anyway.

On the night of 21 June, they drove down quiet Blomberg Road in Olympia, past a dairy and a few houses. They stashed 5-gallon buckets of fuel around a wooden one-storey building, then lit them with barbecue lighter sticks. A national wildlife research facility, run by the

federal government's Animal and Plant Health Inspection Service, was utterly destroyed. It served as lab and office space for government scientists who study animals that eat trees and invent ways to keep them from doing so. At the time of the arson, the lab was working on chemical repellents for beavers and a long-term study of what induces bears to dine on Douglas fir bark.

The attack wouldn't be the last, or the largest, on a scientific facility. Seven years later, in December 2005, the US government indicted 11 men and women on charges of conspiracy and arson for 17 fire-bombings whose results included \$12 million in damage at a ski resort in Colorado as well as the destruction of the Olympia lab. The 11 had come together loosely under the banner of the Earth Liberation Front (ELF) and the Animal Liberation Front (ALF). Their trial, for those not dead, fled or pled out, will be held in Eugene.

For a certain segment of the radical environmental movement, science is seen as both useful and oppressive. Activists sometimes use scientific findings to support their arguments

but often view science, and especially technology, with deep distrust. To them, science is a string of proclamations issued by men in white coats and does not respect their deeply felt connection to individual places, animals and trees. Researchers, particularly environmental scientists, should be aware of such thoughts — not least because to some of these activists, science is a legitimate target in a war in which the very future of Earth is at stake.

Terror list

Although it is impossible to know how large such pockets of radicalism are, the FBI considers them a serious threat. Spokesman Bill Carter says, "From January 1990 to June 2004, animal and environmental rights extremists have claimed credit for more than 1,200 criminal incidents, resulting in millions of dollars in damage and monetary loss." The bureau has officially labelled both groups terrorists — a word freighted with emotion, rhetoric and law-enforcement response.

Dale Nolte, who ran the Olympia facility at the time of the arson, remembers an ELF

F. GRAVES/REUTERS

communiqué claiming credit. “They said they had targeted us because of work we had done with beaver traps and cougars,” says Nolte, now mammals programme manager at the National Wildlife Research Center headquarters in Fort Collins, Colorado. “I don’t know where they got their information, because we had never worked with cougars and we were working with repellents, not traps.”

Nolte has never understood the arsonists’ motives. He lost a personal library in the fire, as well as several crates’ worth of original data and observations about the mountain beaver going back decades. “Our goal was to establish a better understanding of the relationships between the animals and forest, and to enhance the establishment of trees,” he says. “I have a strong sense of protecting the environment.”

To the activists, a lab that served the interests of lumber companies was fair game. Many of them came from the anti-logging community in Eugene.

Among the outsiders

Quiet, pretty and populated with anti-establishment characters, Eugene is a magnet for the kind of activists who end up in the ELF. On a fine day at the downtown Eugene market, one might see the following clues that one has stumbled deep into the heart of a west-coast radical paradise: people bristling with clipboard petitions on a range of left and anarchist issues; others wearing fairy wings; kids in home-made nappies being toted behind bicycles; tie dye, drum circles and the sweet reek of marijuana.

The taxonomy of subcultures here is nearly impenetrable to an outsider. There are anarchists, feminists, Marxists, primitivists, old-school hippies, libertarians and socialists. The groups tend to spend their time debating with each other, more or less ignoring the mainstream.

In June, friends and sympathizers of the Eugene-based eco-prisoners held a benefit at a local bar. The night was mellow, the band good, the beer organic — just the sort of ideological cocoon that can make radical ideas seem pedestrian. Primitivist writer Derrick Jensen gave a lecture on the futility of working within any system to save Earth. For those in the room, his contention that humanity is in “the thrashing endgame of civilization” seemed obvious. A man in a cowboy hat and black overalls, smoking a cigarette on the back porch, said he was ready to give up civilization completely: “Everything, man, the whole gig.”

Jensen, a middle-aged man in a T-shirt reading “Better dead than domesticated”, proved a good rhetorician. Unlike many radical speakers, he was also very funny. But the logic of his arguments sometimes fell apart, and he displayed a strong anti-science streak. He referred to “the myths of science” and said it

doesn’t take a rocket scientist to see that industrial civilization is not sustainable. “In fact,” he continued, “it probably takes everybody but a rocket scientist.”

Jensen’s attitudes are not unusual among ELF-sympathizing eco-activists. These people are deeply sceptical of anything having to do with that entity variously known as ‘the dominant culture’, ‘the white supremacist capitalist imperialist patriarchy’, ‘the man’ or ‘the machine’. Science is part of ‘the machine’, based, in their view, on hard rationalism and cold technology.

Philosophers of science have produced reasonable arguments against science as the pure, objective, single road to truth. But something bad happens when you mix half-baked versions of these critiques with a soul-shaking environmental fervour. What you get is anti-intellectual spasms of violence. This approach is a long way from the earlier tradition of direct action in favour of the environment, pioneered by the group Earth First!

Many scientists were among the first members of that group, formed in 1980 in the US southwest. As Susan Zakin describes them in *Coyotes and Town Dogs*, her 1993 book on the movement, they were boozing, redneck ‘desert rats’ — men who had seen the mesas mined, the Colorado River dammed and the great American West becoming tamed. They stressed ecosystems over scenery, unlike many environmentalists of the time. They never resorted to arson: rather, they dismantled construction equipment, trashed billboards and spiked trees with nails to break chain-saws — but they generally told loggers where they had done this to avoid nasty accidents.

Later in the 1980s, a schism opened between the ecologists in the group and those who were more interested in “monkey-wrenching” — sabotage or other forms of troublemaking. The scientists began to drift off. “By the time we began sitting in redwood trees in northern California, the pagan vegan pacifists outnumbered the biologists ten to one,” says founding member Mike Roselle.

The non-scientists that were left, he says, spawned the ELF. They were radical and committed. “These folks are often the unsung heroes, the true foot soldiers of the movement,” he says. “The ones in jail were among our best, some of the most skilled and committed members of our movement. Their incarceration is a tragic loss for all of us.”

Earth First! co-founder Dave Foreman doesn’t quite agree. “We wanted to use the science of ecology to guide us,” he says. “The current group of people I don’t consider conservationists, but part of the international anarchist animal-rights movement. It becomes sort

of an inarticulate yowl against the establishment — revolution for the hell of it.”

The yowl often starts young. Chelsea Gerlach, one of the recent indictees, was a 16-year-old high-school student in Eugene when she was arrested at an Earth First! logging blockade in Idaho, according to her support website. At school, she ran an environmental club, and, according to Seattle’s alternative weekly, *The Stranger*, was quoted in the yearbook as saying, “Our generation was born to save the Earth. If

we wait until we’re out of school it might be too late.” She was 23 when, the FBI indictment alleges, she helped set fire to the Jefferson Poplar Farms in Clatskanie, Oregon, spray-painting “ELF” and “You cannot control what is wild”.

That was 21 May 2001. For maximum effect, the FBI says, on the same night three others — William Rodgers, Stanislas Meyerhoff and Briana Waters —

torched the University of Washington’s centre for urban horticulture, about 230 kilometres away in Seattle. They took time to remove some cages for pet snakes from the building, Merrill Hall, before setting it aflame.

The Merrill fire

The hall housed lab space for botanists and ecologists, and served as a meeting place for Seattle’s horticulturalists. Assistant professor Sarah Reichard, a specialist in rare plants, remembers the day well. “It was a beautiful morning,” she says, until she got the message that her lab was on fire. She rushed to the hall, which sits just down the hill from the main University of Washington campus and is surrounded by demonstration gardens, meadows, and a grove of native trees.

“There was Merrill Hall with flames leaping 30 or 40 feet into the air,” she remembers. The young academic lost everything, including a tissue culture lab where she was growing the highly endangered showy stickweed, a blue-white flowering plant in the forget-me-not family. Losses of specimens, along with irreplaceable books and slides, put her career back at least a year, she estimates.

It wasn’t just the blow to her road to tenure that worries her. That morning, standing by the blazing building, she feared that someone had been killed. “Academic units are not nine-to-five places,” she says. “In fact, it was very unusual that there was no one there at three in the morning. One graduate student told me that the only reason he wasn’t there was because I gave an extension on an assignment.” No one was hurt in the blaze, but the building was a total loss.

Reichard may have been the hardest hit but she wasn’t the target of the arson. “We could see Toby’s office was black — a big black hole,” she says. “Everyone realized immediately that it was not an accident.”

“The ones in jail were among our best, some of the most skilled and committed members of our movement.”

— Mike Roselle

“It doesn’t take a rocket scientist to see that industrial civilization is not sustainable.”

— Derrick Jensen

LEFT: FBI/AP; CENTRE: LANE COUNTY ADULT CORRECTIONS/AP

Toby is Toby Bradshaw, a plant geneticist who at the time ran the Poplar Molecular Genetics Cooperative, a group working on finding genes in hybrid poplars that code for traits useful in a crop tree — fast-growing, disease-resistant and straight. The team used traditional breeding techniques, ultimately aiming to make productive tree farms more attractive than logging big trees from old-

growth forests. Somehow, however, the radical greens got the idea that Bradshaw was genetically engineering trees.

Bradshaw says he has never done so, although his colleagues have. He'd been a target before. Someone had tried to overturn his potted seedlings during the protest against the World Trade Organization meeting in Seattle in 1999. And two weeks after the Merrill Hall fire, a communiqué was issued through Craig Rosebraugh, a former spokesman for the ELF who has not been indicted. It reads, in part, "Bradshaw, the driving force in G.E. tree research, continues to unleash mutant genes into the environment that is certain to cause irreversible harm to forest ecosystems. As long as universities continue to pursue this reckless 'science', they run the risk of suffering severe losses. Our message remains clear, we are determined to stop genetic engineering."

Bradshaw likes to joke that the fire actually helped his career. He got tenure and a new lab shortly afterwards, and has since moved on to unrelated work. But the attitudes behind the crime trouble him. "As social commentary,



Josephine Sunshine Overaker: fugitive.



Kevin Tubbs: pleaded guilty to arson and conspiracy.



William Rodgers: killed himself in jail.

FBI director Robert Mueller gave a press conference on the indictments in Washington DC. "Terrorism is terrorism, no matter the motive," he said. "The FBI becomes involved, as it did in this case, only when volatile talk crosses the line into violence and criminal activity." Terrorism, however, is not defined as a crime in the United States; the group was charged with arson and associated crimes.

these kinds of arsons are ineffective because they are so misguided," he says. "Science at its heart is a rational enterprise and, at its heart, this kind of terror tactics with fire-bombing is an irrational enterprise."

The crackdown

The FBI last interviewed Bradshaw at the time of the arson, and he was surprised last December when the arrests were announced. In fact, the FBI had spent years putting together information for the indictment; the most recent of the 17 arsons listed in the charges dated to October 2001, at a Bureau of Land Management wild-horse facility in Litchfield, California. The small group that had acted under the ELF banner seems to have more or less broken up after that, scattering to Virginia, Arizona, and around the Pacific Northwest. Reportedly, the FBI laid the groundwork for the charges by getting activist and heavy-metal guitarist Jake Ferguson to call his old pals for some nostalgic, and wire-tapped, conversations.

The bust was important enough to bring out the top law-enforcement brass. On 20 January,

Most of the indictees are now in jails in Oregon, with some under house arrest or out on bail. Lauren Regan, head of the Civil Liberties Defense Center in Eugene, calls the arrests the 'green scare', a play on the 'red scare' of the 1950s in which US citizens with communist ties were persecuted.

Regan says that many of the indictees don't deserve the sentences they are facing — life several times over for each. "A lot of these people were at the time very young," she points out. "You are easily swayed, you've got a lot of passion. You hold a radio while some genetically modified trees are burned down. Chances are that they were not thinking, in that pre-9/11 time, that they were looking at life in jail."

Repentance and escape

Many of those arrested have since pleaded guilty and are cooperating with the authorities. Gerlach made a public statement at the time of her plea, saying, "These acts were motivated by a deep sense of despair and anger at the deteriorating state of the global environment and the escalating inequities within society. But I realized years ago that this was not an effective or appropriate way to effect positive change."

Those outside jail have reacted harshly to those cooperating with the government. In Eugene, Jensen began his speech with a message: "What you are doing is wrong, and I plan on seeing you brought to justice. And fuck you." Ferguson, who never faced charges, is known in some activist circles as 'Jake the Snake'.

A few of the indictees are still at large. Their wanted posters reveal a bit about them. Justine Overaker's, for example, paints a portrait of her as a seeker; it mentions that she may seek work as a "firefighter, a midwife, a sheep tender or a masseuse" and that she can speak Spanish and has been known to use narcotics. She has a tattoo of a bird across her back.

Rodgers, also known as Avalon, was arrested at his bookstore in Prescott, Arizona. He suffocated himself to death with a plastic bag in his jail cell. He left a note that said, "Human cultures have been waging war against the Earth for millennia. I chose to fight on the side of bears, mountain lions, skunks, bats, saguaros, cliff rose and all things wild. I am just the most recent casualty in that war. But tonight I have made a jail break — I am returning home, to



The activists' enemy at work: a loader stacking logs in the Cascade Mountains near Eugene, Oregon.

W. MORGAN/CORBIS



Signature style: a message claiming responsibility for setting fire to an apartment complex being built in Sutter Creek, near Sacramento, California.

the Earth, to the place of my origins."

At 41, Rodgers was the oldest of the indictees and by some reports the ringleader. His lover, Katie Rose Nelson, says, "Life mattered to him — and that meant all life". She and Rodgers both felt the environmental cause was urgent — too urgent to spend time on research. "In our hearts we can all see what is happening around us," she says.

With most indictees not talking, missing or dead, they can't explain the motivations and justifications for their alleged crimes. But a rare glimpse of three of the activists, including Ferguson and Rodgers, can be seen in the 1999 documentary film *Pickaxe*, which chronicles an 11-month battle in the mid-1990s to keep an area in Oregon called Warner Creek from being logged. Here, they engage in legal protests, such as hunger strikes, and the kind of non-violent illegal protest that many condone, such as blockading the road into the area.

And they won. At least that one patch of forest is still there, old and moist and green. So why were these same people drawn to anonymous arson attacks late at night — and why target scientists?

"Science does not have a lock on truth," says David Agranoff, an animal-rights and veganism activist in San Diego, California, who knows some of the defendants. He says he is sure they considered their actions carefully: "I would guess they had a pretty good reason." Rosebraugh, the former ELF spokesman, notes

that each would have had their own motives. "For all the people involved," he says, "you would probably give a different answer."

The 'pure ones'

Donna Haraway, who studies feminism and science at the University of California, Santa Cruz, met some of the radical-activist set in Eugene when she went to speak there in 2001. Because Haraway, author of *Simians, Cyborgs and Women: The Reinvention of Nature*, embraces technology as an agent for positive change, the group saw her as an enemy. Their protest included flyers that she found distinctly threatening. "At least one felt that the rape of nature justified the rape of anyone who supported it," she says. "These people don't go for complexity; they believe that they are the only pure ones who can defend nature."

The ELF indictees are by no means the last

of the self-anointed environmental 'pure ones'. Jeffrey Luers of Eugene is appealing a 22-year sentence for setting fire to sports utility vehicles as a protest. Activist Trey Arrow is fighting extradition from Canada to the United States, where he faces charges of having burned logging and mining trucks. In January, three people were arrested in the parking lot of a Kmart while buying supplies, allegedly to bomb a forest genetics lab in Placerville, California: they are charged with conspiracy, and two have pleaded guilty, although many activists blame an FBI agent provocateur for the plot. In Canada, a group of eco-saboteurs are currently on a fire-bombing tour of construction sites across Ontario. By its nature, the movement is decentralized, non-hierarchical and open. Anyone can be an environmental arsonist.

From jail, Luers publishes a magazine called *Heartcheck*, with an estimated circulation of 1,000. For the latest issue, he has penned an essay called 'Time's up' in which he summarizes recent facts and figures on species extinction in various ecosystems, glacier melt rates in Alaska and carbon dioxide levels. He ends: "I could go on, but you get the point. Or do you? We are not running out of time, we are out of time! We have to act now just so it doesn't get any worse. Smash it, break it, block it, lock down to it. I don't care what you do or how you do it. Just stop it. Get out there and stop it."

Emma Marris is a Washington correspondent for *Nature*.



"This was not an effective or appropriate way to effect positive change."

— Chelsea Gerlach

DRIVEN TO MARKET

We're selfish and rational — that's what classical economics says. But play parlour games with brain scanners and you'll find we're pulled in different directions when it comes to money. **Jonah Lehrer** reports.

Read Montague spent the summer of 2003 thinking about soft drinks. His teenage daughter was working as an intern in his lab at Baylor College of Medicine in Houston, Texas, and Montague, a neuroscientist, wanted to find an experiment that she could "wrap her head around". After much deliberation, he came up with the perfect research topic: recreating the Pepsi Challenge. In a brain scanner¹.

Pepsi launched this advertisement, one of the most famous of all time, in the early 1980s. Television ads showed people on the street being asked to sip cola blindly from two different glasses. Not surprisingly, the ads featured Coca-Cola aficionados who, much to their astonishment, found they preferred the taste of Pepsi.

But if Pepsi really tasted better, Montague wondered, then why would Coke still be more popular? When we are standing in the supermarket, faced by cans of Coke and Pepsi, what is happening inside our brains?

Montague is at the cutting edge of a new scientific field known as neuroeconomics, which uses the experimental techniques of neuroscience to understand how the brain makes economic decisions. Biology, of course, has long been used to explain human nature; evolutionary biology seeks the causes of behaviour in terms of its fitness benefit; and cognitive psychology aims to model decision-making. Neuroeconomics is different: it seeks to understand the most immediate causes of economic choices by seeing how the brain makes them. By studying the brain at work, neuroeconomists hope to resolve well known anomalies such as why stock markets are sometimes gripped by 'irrational' exuberance, or why people rack up huge credit-card debts to buy things they don't need.

The stubborn persistence of these perplexing phenomena defies classical economics, founded as it is on two assumptions about

human nature: that we are rational and that we are selfish. When confronted with a variety of options, traditional economists expect us to evaluate the possibilities (rationality) and choose whichever best matches our personal preferences (selfishness). Their mathematical models require this predictable behaviour. What the eighteenth-century economist Adam Smith called the "invisible hand" of the marketplace is just the collective result of lots of reasonable people going about the business of trying to maximize their own advantage. Such pure rationality is disconcertingly rare, however. Neuroeconomists want to explain why, and their research promises to affect everything from what cola we drink to how we save for retirement.

Fair play

The story of neuroeconomics began in the early 1980s with a parlour game, the ultimatum game, devised to investigate economic behaviour². The rules are simple. An experimenter puts two people together and hands one of them \$10. This person (the 'proposer') has to offer some of the money to the other. The second person (the 'responder') can either accept the offer, in which case both players pocket their respective shares, or reject it, in which case the \$10 is taken from the proposer, and both players walk away empty-handed.

According to the predictions of classical economics, the game should always generate the same outcome. The proposer should offer the responder a minimal amount, \$1, and the responder should always accept it. After all, \$1 is better than nothing, and a rejection leaves both players worse off. If the ultimatum game played out this way, it would be a clear demonstration of our rational self-interest.

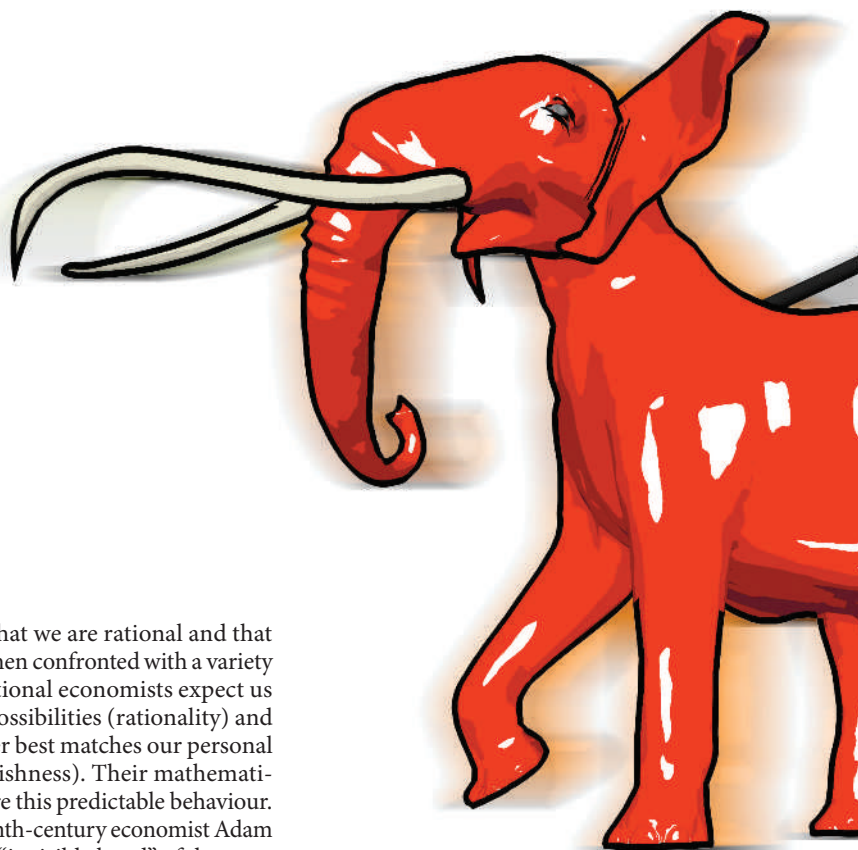
When the game is played, however, that doesn't happen. Instead of taking a small profit, responders typically reject any offer that they think unfair. Proposers tend to anticipate this 'irrational' rejection and, instead of offering a minimal amount, typically propose around \$4. This isn't a cultural prejudice: people around the world play the ultimatum game the same way. The only ones who obey the expectations of classical economics are autistic adults: because they don't take the feelings of the other player into account, they typically offer the minimum amount.

Why would someone make the seemingly

irrational decision to reject free money? Evolutionary game theory provides some insight. In the real world, losing out on the money in the short term could mean getting a social benefit in the long term: a reputation for not being a pushover, for one thing³. Players in the ultimatum game, however, don't have to worry about this: they play each other only once and

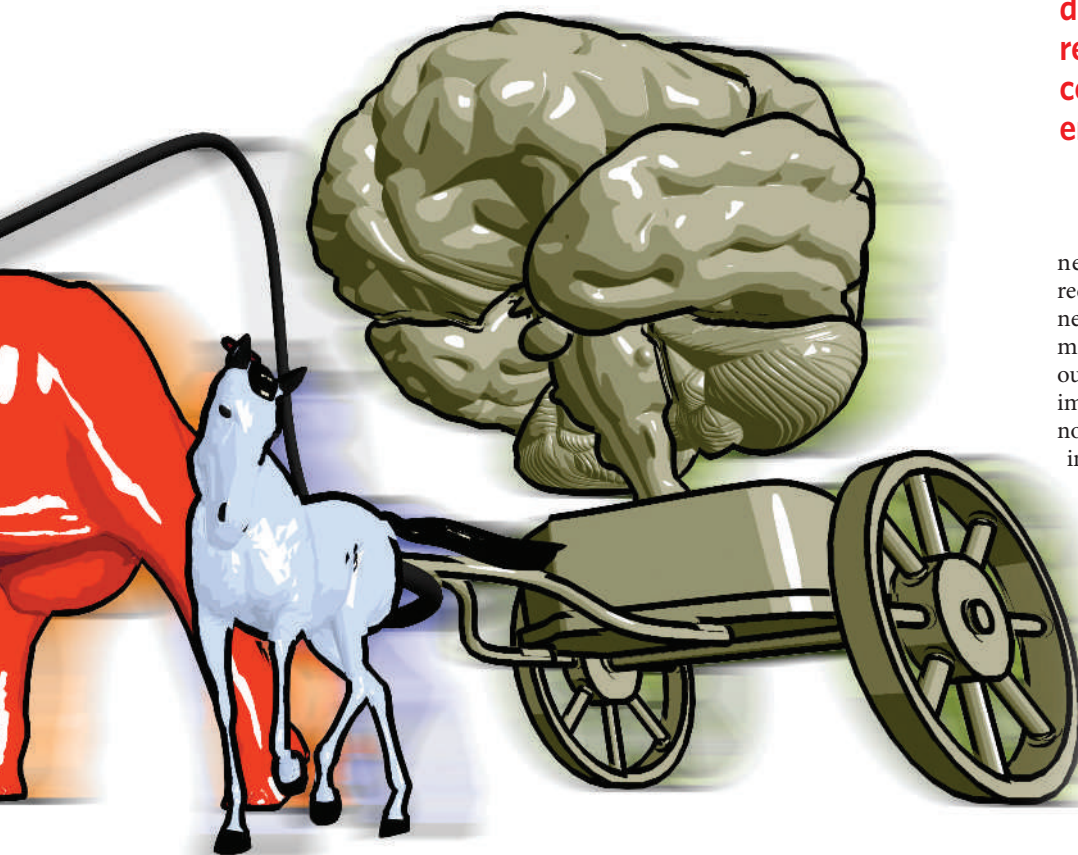
have no information on their partner's reputation. But fairness still trumps reason. So what is going on inside the brain?

In 2003, Alan Sanfey, Jonathan Cohen and their colleagues at Princeton University, New Jersey, used a technique called functional magnetic resonance imaging, or fMRI, to look



"Our emotions are not just a negative impulse that gets in the way of our rationality. They are much more integrated."

— Paul Glimcher



"The mind is a charioteer driving twin horses of reason and emotion. Except cognition is a smart pony, and emotion an elephant"

— Colin Camerer and George Loewenstein

neuroeconomics will require this sort of reductionism "to construct a general theory of neural decision-making". Other neuroeconomists counter that, although fMRI has serious limitations — Camerer admits it's a "very imperfect tool" with some "serious signal-to-noise problems" — it remains useful for giving insight into what the brain is doing.

Many neuroeconomists, however, judge their field by what it adds to economic theory rather than by its neurological precision. In these terms, some argue that the dual-process model has already been a success because it can explain behaviour that economists have not been able to. According to Andrew Lo, director of the Massachusetts Institute of Technology laboratory for financial engineering, "Economics has hit the wall. It has explained about as much as it can with the tools it has. There are too many inconsistencies between theory and data."

Thinking ahead

One anomaly that continues to confound economists is humanity's often irrational approach to the future. Instead of saving money for retirement, people tend to impulsively splurge on the present. Neuroeconomists are beginning to understand the neural roots of this behaviour. In a 2004 brain-imaging experiment led by Samuel McClure of Princeton, people were asked whether they wanted a low-value Amazon gift voucher now or a higher-value voucher in two to four weeks⁵. McClure wanted to test a specific assumption of classical economics: the idea that we apply the same calculus to the future and the present. If that were true, then the same brain regions should become active whether we are thinking about the results of economic decisions in the future or in the present.

This isn't what McClure found. When his subjects contemplated receiving gift vouchers in the future, brain areas associated with rationality (such as the prefrontal cortex) became active. These cortical regions seemed to urge people to resist temptation and wait for the more valuable vouchers.

On the other hand, when people started thinking about getting a gift voucher right away, brain areas associated with emotion — the midbrain dopamine system, for instance — were also turned on. By manipulating the value of vouchers in each situation, the researchers

inside the minds of people playing the ultimatum game⁴. fMRI scans highlight areas of the brain that are more metabolically active than others and are therefore thought to have increased neuronal activity. The team found that 'unfair' offers led to brain areas such as the anterior insula, associated with strong, negative emotions such as disgust and pain, becoming more active. At the same time, there was activation of brain areas associated with information processing and long-term planning, such as the dorsolateral prefrontal cortex (DLPFC).

In two minds

When subjects were struggling to decide whether or not to reject an unfair offer, those whose anterior insula showed more activity than their DLPFC tended to reject unfair offers, whereas those whose brains exhibited the opposite pattern tended to accept them. This, says the team, suggests that competition between these areas influences decision-making — and emotions usually win. "The platonic metaphor of the mind as a charioteer driving twin horses of reason and emotion is on the right track," wrote the neuroeconomists Colin Camerer of the California Institute of Technology in Pasadena and George Loewenstein of Carnegie Mellon University in Pittsburgh, Pennsylvania, in an unpublished working paper. "Except that cognition is a smart pony, and emotion a big elephant."

This interpretation, in which the brain is

capable of both deductive logic and irrational emotion, often simultaneously, is known as the 'dual-process' model and it remains controversial. "Imagine showing this model of cognition to an economist and a neuroscientist," Camerer says. "They'll both think it's wrong, but for opposite reasons. The economist will say it is too complicated: the brain only needs a rational system. The neuroscientist will say it's too simple, and that our brain regions can't be parcelled out into rational and irrational categories. Neuroeconomics is trying to find the middle ground."

Some researchers don't think that middle ground has been found. Paul Glimcher, a neuroeconomist at New York University, warns that the dual-process model is not biologically accurate. "Experiments done on monkeys have never supported this notion of there being two fully independent decision-making systems," Glimcher says. "This doesn't mean that emotions don't exist or that they don't influence our behaviour. What it does suggest is that our emotions are not just a negative impulse that gets in the way of our rationality. They are much more integrated than that."

Rather than focus on brain-imaging research — which Glimcher dismissively calls "spots-on-brain" experiments — his lab records from neurons in specific areas of the cortex while monkeys are making 'economic' decisions, such as trying to maximize a reward of fruit juice. Glimcher argues that rigorous

could compare the levels of activation in the different regions. They discovered that the relative amount of activity was “directly associated with subjects’ choices”. People whose ‘emotional’ brain areas were more active opted for the spoils of immediate gratification.

This discovery has important implications. For starters, it helps explain why people often fail to save enough for their retirement. Because our emotions warp our better judgement, we delay saving. Loewenstein, who collaborated on the McClure paper, thinks that understanding how we make such decisions will help us develop better economic policies: “Our emotions are like programs that evolved to solve problems in our distant past. They are not necessarily well suited to modern life. It’s important to know how they lead us astray so that we can design incentives and programmes to help compensate for our irrational biases.”

What might a savings scheme informed by neuroeconomics look like? In March 2004, behavioural economist Richard Thaler of the University of Chicago, Illinois, testified before the Senate on ways to increase the national savings rate — US consumers currently have a savings rate close to zero⁶. His plan was simple: rather than asking people if they want to start saving right away, companies should ask people if they want to opt into a savings plan that begins in a few months’ time.

This allows people to make decisions about the future without contemplating the present, bypassing our irrational emotions. McClure’s brain research suggests that should be a smart approach, and indeed trial studies of Thaler’s plan have been a resounding success: after three years, average savings rates jumped from 3.5% to 13.6%.

Credibility gap

Many experts remain sceptical of neuroeconomics, however. The Princeton economists Faruk Gul and Wolfgang Pesendorfer think it is based on a faulty premise. They argue that economic models should be judged by their success at explaining phenomena such as inflation or unemployment, not by “psychological realism”.

In an attempt to close that credibility gap, neuroeconomists are trying to bring their experiments closer to the decision-making models of microeconomics, which studies individual behaviour. If they can’t scan people’s brains in the real world, they can at least bring a little bit of the real world into the lab. Take, for example, Montague’s Pepsi Challenge experiment. Rather than playing parlour games in an fMRI machine, he monitored people’s brain activity as they swallowed sips of soda¹.

When the Coke and Pepsi were offered unlabelled, people showed no measurable preference for either brand. Most of the time, they couldn’t even tell them apart. Montague’s



Read Montague puts his money where his mouth is in his rigorous version of the Pepsi Challenge.

second observation was telling: people strongly preferred drinks that were labelled as Coke, no matter what cola was actually delivered through the tubes. Brand trumped taste.

The fMRI scans revealed that when the drinks were offered unlabelled, the ventromedial prefrontal cortex (VMPFC) became active. This makes sense, because the VMPFC is involved with the processing of appetitive rewards such as sugary drinks. However, when the subjects drank a cola with a Coke label, more brain areas were turned on. The hippocampus and midbrain all reacted strongly to the red curvise of Coke but not to the blue Pepsi logo. This happened even

when subjects were given Pepsi with a Coke label. Montague notes that the brain regions triggered by Coke have all been implicated in ‘affective’ — emotional — influences on behaviour. Brand power exerts a more powerful force over our emotions and decisions than we might like to think. “Advertising is a deeply biological game,” Montague says. “The idea of Coke clearly affects our judgement.”

Trust me

People rarely make economic decisions in isolation like this: most involve interacting with others. So Camerer, Montague and Steve Quartz, also at the California Institute of Technology, decided to link their fMRI machines together and monitor different brains simultaneously⁷. This allowed them to measure brain activity during social interactions, as subjects were learning to trust each other.

They invented a simple trust game in which an ‘investor’ has the option of entrusting money to a ‘trustee’. Invested money gets tripled, and the trustee can then keep it all, or give some or all of it back to the investor. Because the game is typically played for ten rounds, however, with the investor receiving more cash to play with

each round, each player has a selfish incentive to trust each other.

The researchers discovered that increased activity in the caudate nucleus — a region involved in the brain’s reward pathway — of the trustee was directly correlated with the trustworthiness of the investor’s behaviour. Furthermore, this activity appeared much more quickly in later rounds of the game, indicating that the trustees were forming an opinion of their partners.

Why the caudate nucleus? The neuroeconomists speculate that a decision to trust someone else depends upon our expectation of getting a reward in the end. The caudate nucleus gets excited by the anticipation of material pleasures such as food, drugs and money, so it makes sense that it would also measure the rewards of our social interactions. Trust is therefore an admirable trait with selfish origins.

Neuroeconomists are bullish about their own potential. “Economics provides us with all sorts of elegant theories, and neuroscience gives us a new way of testing their predictions,” says Loewenstein. “All we need is time.” Daniel Kahneman, a psychologist at Princeton, agrees. “These researchers have the chance to come up with a general theory of decision-making that is both biologically and behaviourally accurate. They have the necessary experimental tools and are asking the right kind of questions. Now they just have to find some answers.” ■

Jonah Lehrer is a science writer based in Boston, Massachusetts.

1. McClure, S. M. *et al.* *Neuron* **44**, 379–387 (2004).
2. Güth, W., Schmittberger, R. & Schwartz, B. *J. Econ. Behav. Organ.* **3**, 367–388. (1982).
3. Nowak, M. A., Page, K. M. & Sigmund, K. *Science* **289**, 1773–1775 (2000).
4. Sanfey, A. G., Rilling, J. K., Aronson, J. A., Nystrom, L. E. & Cohen, J. D. *Science* **300**, 1755–1758 (2003).
5. McClure, S. M., Laibson, D. I., Loewenstein, G. & Cohen, J. D. *Science* **306**, 503–507 (2004).
6. Thaler, R. Helping Americans save (US Senate, 2004); available online at http://jec.senate.gov/_files/ThalerTestimony03102004.pdf (2004).
7. King-Casas, B. *et al.* *Science* **308**, 78–83 (2005).

“Advertising is a deeply biological game. The idea of Coke clearly affects our judgement.” — Read Montague

A NOVEL REALITY

Can an advertising executive write an accurate thriller about science?

Britta Danger talks to a German author who thinks he has pulled it off.

Marine biologist Sigur Johanson is wary when Tina Lund, an oil-company scientist, visits his remote Norwegian home. With mixed feelings — he has long been secretly attracted to her — he agrees to help her employer with a serious problem: swarms of *Hesiocaeca methanicola*, or ice worms, at potential North Sea oil-drilling sites. Other scientists join in too, as the invasion turns out to be related to the mysterious breakdown of methane hydrate reservoirs on the sea bed off Europe.

But worms are not the only creatures swarming. Mussels, sea wasps, and Portuguese men-of-war also gather in untold numbers, killing seafarers and beach tourists. A tsunami hits northern Europe. Death and destruction is everywhere. As crisis becomes global catastrophe, the US government — represented by the ambitious, single-minded General Judith Li — takes over the worldwide effort to unravel the mystery.

If *The Swarm* sounds destined for the silver screen, it is: Hollywood actor and producer Uma Thurman has already snapped up the film rights to the best-selling German thriller, now hitting the shelves in an English-language translation. The book's author, Frank Schätzing, says he based his plot on a dream. "Sea life was flocking together, threatening us," he remembers. On waking, he began to wonder: "What would happen if...?"

Yet much of *The Swarm* is based in reality — indeed, it contains a level of scientific verisimilitude unusual in any novel, let alone an airport thriller. Schätzing, a self-assured, well-coiffed 49-year-old, seems at first glance an unlikely figure to have authored such a book. He is an advertising executive in Cologne, and none of his previous five novels dealt with science. Yet readers have found themselves absorbed not only in the tensions and romances of a tough mystery, but also in the details of up-to-the-minute research in fields such as neurocomputing, seafloor oceanography and cell signalling. With no formal scientific education, Schätzing has managed to cover a swathe of scientific territory without significant error.

Which is not to say that the novel is a truly realistic portrayal of the oceanographic world. The antagonist, after all, is pure



Total immersion: to write *The Swarm*, Schätzing delved deep into seafloor oceanography and other cutting-edge science.

fantasy: "I created an environment as real as possible and added only one fictional element — a deep-sea, non-human intelligence crucial to the plot," says Schätzing. He has Johanson dub this intelligence Yrr by striking a computer keyboard three times at random.

Schätzing developed the plot over three years of devouring popular science books, research reviews and internet sources. His studies led him to the dozen or so researchers with whom he hammered out the details of how the science could be made to serve the plot without becoming distorted. "With them, I tried to see how far I could stretch my ideas," he says. "We developed methods together, for example the way the Yrr communicated" — through pheromones in the water.

Schätzing visited scientists across Germany and further afield, including in Vancouver, Canada. He developed a particularly close relationship with those at the Leibniz Institute of Marine Sciences at the University of Kiel in Germany. There he spent hours discussing his ideas with researchers including Erwin Suess, a prominent methane-hydrate researcher; marine biologist Heiko Sahling; and marine geologist Gerhard Bohrmann. All showed up in the novel under their own names, to their initial consternation. But once they had

read the book, none of them minded.

"I was worried when I started reading, but in the end I found myself nicely characterized," says Bohrmann, who has plenty of adventures in real life on research vessels in stormy seas — albeit not on the scale of his fictional namesake, who escapes a vicious shark attack while trying to save humanity from the ice worms. Schätzing smiles, recalling how he originally intended Bohrmann to play a tiny role. "But then he ended up as the Bruce Willis of marine science."

Bohrmann says that Schätzing had done his research well before coming to Kiel for an interview: "He already knew everything about gas hydrates." Sahling was also impressed: "His questions were very intelligent and he was a good listener — he adopted much of what we said word-for-word in his novel." Suess was less charmed, however, remarking that Schätzing claimed as his own ideas that were generated in discussion with colleagues. Science communicator Thomas Orthmann was similarly unimpressed and tried to sue Schätzing for plagiarising from his website. His case was unsuccessful.

Schätzing did not speak to all the scientists he characterized. Ryo Matsumoto, a gas-hydrate expert from the University of Tokyo, was tipped off about his role by his German colleagues. Curious yet nervous, Matsumoto had to wait two years for the English translation to be published to find out how he appeared in the novel. But he was pleased with his modest role as the scientist who confirms that ice worms have reached the Japanese Pacific. "I am surprised at his knowledge of biogeosciences," he says.

Despite his interest, Matsumoto has yet to finish the book. At more than 900 pages, even avid supporters admit that it is simply too long. Descriptions of the science — often disguised as discussions between researchers, or as lectures — can run on for pages. This is both the fascination of the book and its literary weakness, as it drastically slows the plot.

As a novel *The Swarm* may be unlikely to repeat its phenomenal German success out in the English-speaking world. But its future in Hollywood looks brighter — it is, after all, a thriller packed with technology, danger, spectacle, romance and a watery hint of apocalypse. The prospect certainly has Schätzing in dream mode again: George Clooney and Lucy Liu in the main roles — and a cameo for himself, Hitchcock-style.

Britta Danger is a former intern at *Nature's* Munich office.

"His questions were very intelligent, and he adopted much of what we said word-for-word in the novel."

— Heiko Sahling

Speaking different languages on biodiversity

SIR — Interdisciplinary conferences such as BioEcon (Biodiversity and Economics for Conservation), whose eighth meeting “Economic Analysis of Ecology and Biodiversity” has just finished, provide much-needed forums to cultivate the scientific collaborations on which mutual understanding depends. Economic context, in particular, is essential for understanding and reducing the ongoing pressures on biodiversity, as well as for developing successful, cost-effective and equitable conservation actions.

As conservation biologists at BioEcon, we were not expecting to be familiar with the intricacies of economic models. But we had failed to anticipate the differences in vocabulary in what we presumed to be common ground between our disciplines. Presentations on ‘agricultural biodiversity’, for example, turned out to be about crop and livestock genetic diversity, rather than, as conservation biologists would expect, on the diversity of wildlife associated with agricultural fields. In another example, the term ‘statistical dichotomy’ was used in place of our familiar ‘binary variables’.

We were thus amused to read in “100 years ago” (*Nature* **442**, 989; 2006) an accurate description of communication difficulties between scientific disciplines: “progress is not a little hampered by the fact that chemists and physicists cannot wander through the museums of nature in complete sympathy with one another... a confusion of language has arisen which keeps us apart”. The world’s threatened biodiversity cannot wait another 100 years for biologists and economists to overcome their language barriers.

R. M. Ewers*†, **A. S. L. Rodrigues†**

*Institute of Zoology, Zoological Society of London, Regent’s Park, London NW1 4RY, UK

†Conservation Science Group, Department of Zoology, University of Cambridge, Cambridge CB2 3EJ, UK

Bench-to-bedside solution to funding problems

SIR — Recent budget allocations by the US National Institutes of Health (NIH) have aroused serious concern among researchers (see, for example, A. R. Marks *J. Clin. Invest.* **116**, 844; 2006). The problem seems to stem from the division between those who are funded by the NIH and those who direct it. A great proportion of the former are doctors of research, whereas many of the latter are doctors of medicine. As its website shows, every director of the organization has held an MD degree. Unlike those seeking grants, only

4 out of 14 current advisory-committee members hold a PhD (see links from www.nih.gov/about/director/index.htm). Hence, disagreement is not surprising during a funding freeze.

This problem could, in part, be overcome by bridging the gap between medical and PhD students — a gap that is already narrowing, as noted in your News Feature “Them and us no longer” (*Nature* **439**, 779–780; 2006). The emergence of ‘bench-to-bedside’ interdisciplinary research requires and promotes mutual understanding between physicians and scientists.

Yet the road ahead seems long, and many prejudices still exist. It is crucial to introduce new programmes aimed at acclimatizing medical students to the research process and PhD students to clinical practice. As a graduate student, I find this lacking in most institutions. The exception is a limited MD–PhD programme, only available to a handful of students. This concept should be expanded into an integral part of medical education, one that allows for greater understanding and interaction between graduate students and medical students.

Samuel F. Bakhom

Molecular and cellular biology programme, Dartmouth Medical School, Hanover, New Hampshire 03755, USA

Learning from painful experience of disaster

SIR — Your Editorial about preparing for natural disasters (“State of readiness” *Nature* **442**, 847–848; 2006) provides some good advice. As someone whose lab was out of commission for two years after the Northridge earthquake of 1994, I would like to suggest some more precautions, and lessons learned, that might be helpful to other readers.

First and most important, have at least two sets of all important data, one kept in the lab and the other at the home of the project director. It is best to have both electronic and paper copies of all important data. All expensive equipment should be kept off the floor, if possible, to avoid ground-water damage and should be covered with plastic sheeting when not in use to avoid damage from water coming through the ceiling. Expensive equipment should be secured so that it will not fall off counters — but security devices should be easily removable in case items have to be moved quickly. Finally, chemicals should be stored where they can not easily be jostled to the floor, using cabinet locks and shelf lips. These few tips could help save your lab in the event of a disaster.

Steven B. Oppenheimer

Center for Cancer and Developmental Biology, California State University Northridge, Northridge, California 91330-8303, USA

Hoyle’s observations were right on the ball

SIR — In his In Retrospect article “Out of the darkness” (*Nature* **442**, 986; 2006), Jay M. Pasachoff notes the impact of Fred Hoyle’s novel *The Black Cloud* (first published in 1957) and its contemporary relevance. We would like to add a further aspect: the contribution of two ideas in Hoyle’s book to vision research and to our understanding of the perceptual guidance of action.

The characters in the book discover an ominous black cloud that appears to be heading towards Earth. Will the cloud hit Earth and, if so, when? The first question is solved when the characters examine the relative speed at which the cloud is translating across the night sky to the rate at which it is looming, or seeming to get larger. The second question is tackled with a bit of impromptu algebra in which the time until impact is calculated from the ratio of the current size of the cloud to its rate of change. A mathematical derivation of the formula is provided.

A footballer wishing to head an approaching ball needs to know where the ball is going relative to the head, and when it will hit or pass the head. The player could estimate the trajectory of the ball from knowledge of its position and velocity. However, David Lee realized in the 1970s that the brain can use the ratio of size to its rate of change, previously identified by Hoyle, to estimate the imminence of arrival. David Regan realized soon afterwards that the brain can use the ratio of lateral speed to looming rate to calculate where an object is travelling. These elegant solutions bypass the need to know position and velocity, so that the two quantities of interest can be estimated directly. This human ability is important for the characters in Hoyle’s story because the position and velocity of the cloud are unknown.

An ability to estimate these quantities is of use not just for heading a ball, but also when trying to cross a street full of cars, return a tennis serve or pick up a cup of coffee while rushing past your desk. Since the early work of Lee and Regan, a considerable amount of research in areas including psychophysics, motor action, neurophysiology and computational modelling has followed (see D. Regan and R. Gray *Trends in Cognitive Sciences* **4**, 99–107; 2000). The whole body of work that exists today can be traced back to a casual footnote and a couple of sketches in a science-fiction novel.

Simon K. Rushton*, **Rob Gray†**

*School of Psychology, Cardiff University, Tower Building, Park Place, Cardiff CF10 3AT, UK

†Department of Applied Psychology, Arizona State University, 7001 E Williams Field Road, 340J Sutton, Mesa, Arizona 85212, USA

BOOKS & ARTS

Unburdened by proof

String theorists are setting a worrying trend by downplaying the need for experimental evidence.

The Trouble With Physics: The Rise of String Theory, the Fall of a Science, and What Comes Next

by Lee Smolin

Houghton Mifflin: 2006. 416 pp. \$26

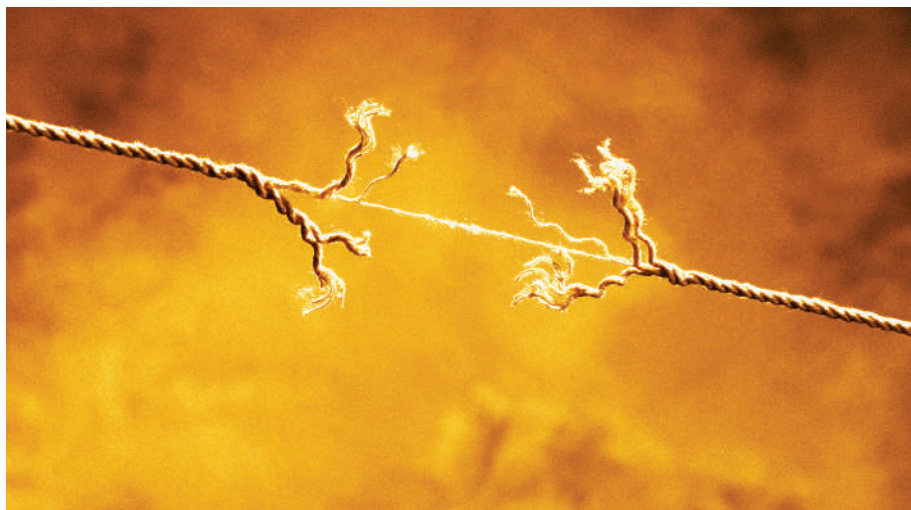
George Ellis

String theory — the proposal that physics is based on two-dimensional entities ('strings') vibrating in ten-dimensional space-time — is viewed by many physicists as the most promising approach to unifying all the fundamental forces of nature. It could, they claim, provide the ultimate 'theory of everything'. But they are seemingly facing a dead-end in their attempts to turn string theory into an experimentally verifiable, fundamental theory of all physics.

There are several problems. The mathematics is so complex that realistic solutions are hard to achieve. The energies involved in testing most versions of the theory are beyond those attainable in particle accelerators. And a recently discovered 'landscape' of string theory predicts that the theory can lead to many versions of low-energy physics depending on the circumstances. What's more, because string theory, if true, can explain virtually anything, it cannot be disproved by any local physical experiment. As a result, some influential string theorists are pushing for the traditional evidence-based approach to science to be replaced by one in which the beauty or elegance of a theory largely replaces experimental confirmation as the basis for belief in its veracity.

In *The Trouble with Physics*, Lee Smolin writes that the theory "has failed to make any predictions by which it can be tested, and some of its proponents, rather than admitting that, are seeking leave to change the rules so that their theory will not need to pass the usual tests we impose on scientific ideas". This would leave not just physics but the whole of science facing something of a crisis. If the grounds of proof that have been the basis of scientific success for the past three hundred years are to be changed, our understanding of what is genuine science will be brought into question.

Most of the recent popular books on uniting gravitational theory and quantum physics present a positive view of string theory and its successes. But Smolin's book and Peter Woit's *Not Even Wrong* (Jonathan Cape, 2006) are critical both of string theory itself and of the effect it has had on theoretical physics by



W. CROCKER/GETTY

At breaking point: string theorists' approach to evidence is threatening to split the physics community.

dominating the field in recent decades.

Smolin begins with an excellent presentation of the foundations of fundamental physics, laying the ground for an understanding of the roots of both the present aims of string theory and its problems. He makes the excitement and promise of such a unifying theory come alive. He then carefully examines what the theory can and cannot explain, explores the difficulties in subjecting it to experimental verification, and discusses problems in its structure.

One problem Smolin highlights is that string theory does not take on board one of the main lessons of the general theory of relativity: a properly formulated theory of fundamental physics that includes gravity should not be based on a fixed background space-time. A major problem is that string theory is not yet a well-defined theory, just a broad programme of research: "What we know of string theory consists mostly of approximate results and conjectures." It has been conjectured that these various results are unified in a deeper theory called M-theory, which involves entities called membranes living in 11 space-time dimensions, but the nature of this theory is unknown. Discussing the string-theory landscape, Steven Weinberg said: "It wouldn't hurt in this work if we knew what string theory is."

A practical problem is that all explicitly known string theories disagree with established facts about our world — for example, most specific cases where we can do exact calculations have unbroken supersymmetry, which is

not true of the real world. Celebrated results attained with regard to black-hole entropy do not refer to realistic black holes that might occur in astrophysics. And in terms of actual results, string theory makes no new predictions for experimental observations, even though its proponents are aware of the need for such predictions. String theory's major strength is that it impressively unifies the kinds of particles and forces we already know about, but it cannot, for example, predict the parameters of the standard model of particle physics, one of the main aims of such a unifying theory.

String theory's most important prediction is that there are many more dimensions than determined by everyday physical experiment. Either these dimensions are wrapped up so small that we cannot detect them, or we live trapped on a four-dimensional surface (a 'brane') embedded in a larger 'bulk' space-time, and so are unable to directly experience these higher dimensions. But so far there is not the slightest evidence that the claim of extra dimensions is true. Despite this, many string theorists are dogmatic about them because their mathematical theories demand that they exist, even though there is no evidence that these theories apply to the real Universe in which we live. Simpler theories can explain just as much as string theory, but a unified description would be far preferable. However, until we have proof that a unified theory exists, its existence remains no more than conjecture.

Given these problems, some string theorists

now have the more modest goal of understanding high-energy physics by means of some interesting dualities that enable easy calculations to give results for very hard ones. But this aim has not received the same publicity as that of finding a unified 'theory of everything'.

So, a theory that is unable to produce any testable new predictions has dominated theoretical physics for 30 years. Given the difficulty of the task, there is nothing wrong with this: we can hope it will eventually work out, and it is reasonable to expect such a fundamentally important enterprise to take a long time. More problematic is the attitude of some string theorists that it is not worth investigating alternatives. There is an unquestioning assumption that string theory's claims must be true, even though there is no solid evidence for them, and this is often expressed with considerable arrogance and an attitude that nothing else is worthwhile physics. String theory has dominated appointments to academic positions in theoretical physics for decades, effectively impeding a broader investigation of quantum gravity and the fundamental nature of space-time. It is claimed to be the only game in town, but insofar as the aim is to quantize gravity, there are alternatives, and Smolin briefly outlines some of the more promising ones.

Some of the sociological issues that come into play here are the subject of interesting chapters, based on Smolin's own experiences. A sad aspect is the *ad hominem* attacks made on those who question the theory, including serious thinkers being labelled by the derogatory term 'popperazzi'. This term makes clear how some string theorists regard their views as so overwhelmingly convincing that it is no longer necessary to retain experimental testing as the core of the scientific approach. Smolin crystallizes what many in the physics community feel about these extravagances of string theory.

Those advocating a focus on 'beauty' and 'miracles' when evaluating their theories don't seem to have thought through the implications. The weakening of criteria proposed by some string theorists will, if accepted, open the doors to many other faith-based enterprises that would be only too glad to be viewed as science. In particular, scientific opposition to 'intelligent design' centres on an insistence that for a theory to be scientific it must be testable, observationally or experimentally. Proponents of intelligent design must surely welcome the freedom from evidential constraints that some string theorists are proposing.

What is crucially needed in developing string theory is a serious attempt to engage with the philosophy of science, developing an approach to theory validation that is adequate where insubstantial evidential support has to be supplemented by other principles of inference. So far, this has not been done. ■

George Ellis is in the Department of Mathematics, University of Cape Town, Cape Town 7701, South Africa.

Developing diversity

The Regulatory Genome: Gene Regulatory Networks in Development and Evolution
by Eric H. Davidson

Elsevier: 2006. 304 pp. \$69.95, £43.99

Michael Karin

All living organisms deploy similar evolutionarily conserved mechanisms to generate energy, replicate their genomes, use genetic information and synthesize basic building-blocks for their cells. Yet the myriad shapes and forms of both plants and animals are overwhelming in their variety and extremes. What is even more amazing is that most plants and animals start their life as a single diploid cell (a zygote) created by the union of a sperm and an egg. How these simple cells give rise to such complex creatures with diverse body shapes is a major preoccupation of developmental biologists.

Developmental biology mainly deals with the processes and regulatory mechanisms that guide the conversion of a unicellular zygote into a multicellular organism. The progression of body shapes and forms over evolutionary time is another fascinating topic that occupies students of development and evolution alike. As Eric Davidson convincingly argues in his book *The Regulatory Genome*, animal (and plant) development is driven by a "dynamic progression of regulatory states, defined by the presence and state of activity in the cell nuclei of particular sets of DNA recognizing regulatory proteins (transcription factors), which determine gene expression". An equally important contributor to development is the "genomic apparatus that encodes the interpretation of these regulatory states". Over evolutionary time, "the alteration of body plans is caused by changes in the organization of this core genomic code for developmental gene

regulation". The genomic code that dictates animal development and the genetic apparatus that implements it are the main topics of the book.

Initially a descriptive science, developmental biology has benefited hugely from the revolution in molecular biology of the past 30 years. Indeed, *The Regulatory Genome* is entirely different from Davidson's first book, *Gene Activity in Early Development* (Academic Press, 1968), which preceded molecular cloning and rapid DNA sequencing. However, even in the earlier book, which I remember reading as a graduate student, Davidson championed a quantitative, biochemical and mechanistic approach to the study of developmental biology. His latest book goes many steps beyond that, explaining the basic processes of development and evolution in terms of information processing and computational logic. *The Regulatory Genome* incorporates and integrates many recent advances in understanding gene regulation and genomic organization in a quest for a unified model to explain animal development and the evolution of body shapes and forms.

As a scientist interested in gene regulation and signal transduction, I think the most important and valuable message conveyed by the book is the central role of the DNA elements, the *cis*-regulatory elements and control units, in both development and evolution. Most molecular biologists are occupied as I am with the study of regulatory proteins, either transcription factors or signal transducers that eventually modulate transcription-factor activity. So it is a refreshing and sobering realization that the *cis*-regulatory elements — the genomic DNA units recognized by sequence-specific transcription factors — occupy a more central role in the design of the genetic circuits and networks that control developmental



Variety show: how did animals develop diverse body shapes like this nudibranch and these tunicates?

T. LAMAN/GETTY

processes and mediate evolutionary diversification than the transcription factors themselves. This makes sense because, in most cases, sequence-specific transcription factors or the signalling proteins that modulate their activity are highly conserved among organisms with very different shapes and forms. Yet the constellations of *cis*-regulatory elements, which together make up the *cis*-regulatory control units that dictate the time, place and magnitude of gene expression, are more diverse and seem to evolve more rapidly than the transcription factors that recognize them. Furthermore, such control units also dictate the expression patterns of transcription factors and signalling proteins in time and space, and thereby

determine the exact patterns and repertoires of developmental gene expression.

In general, Davidson does an excellent job of reducing the complexity of different developmental pathways and modes of embryonic development in diverse animal phyla to a set of simplified and logical concepts and principles. He provides excellent illustrations and experimental examples derived from several model organisms: nematode worms, fruitflies, sea urchins, tunicates and vertebrates of different sorts. What is especially attractive about the book are the regulatory networks drawn as simple wiring and computational diagrams. These go a long way towards explaining the basic regulatory logic and engineering principles of

some of the most complex biological phenomena: animal development and the evolution of body forms.

This book should be read by all biologists who want to understand how development and evolution take place and what governs the workings of genomes. I also recommend it to computer scientists and engineers who are interested in the budding field of computational biology, as reading it does not require an extensive background in developmental biology. ■

Michael Karin is in the Department of Pharmacology, University of California, San Diego School of Medicine, La Jolla, California 92093, USA.

Putting DNA on the map

Reconceiving the Gene: Seymour Benzer's Adventures in Phage Genetics

by Frederic Lawrence Holmes

Yale University Press: 2006. 320 pp. \$50

Denis Thieffry

In the 1950s, Seymour Benzer set out on a daunting research programme aimed at resolving the fine structure of the gene. Using bacteriophages and various strains of bacteria, his genetic-mapping enterprise ultimately united genetics and structural chemistry. How did Benzer conceive and achieve this goal? And what were the critical components and influences that led to the success of his enterprise?

In *Reconceiving the Gene*, Larry Holmes addresses these questions in a scrupulous and captivating historical analysis. Making extensive use of Benzer's laboratory notebooks, scientific correspondence, reports and grant proposals, and complementing these with interviews, Holmes carefully retraces Benzer's sinuous investigative path and offers us a well-documented day-to-day analysis of the development of his work. Although Benzer had previously published his own biographical recollection of this period as a chapter in the edited volume *Phage and the Origins of Molecular Biology* (Cold Spring Harbor Laboratory Press, 1966), Holmes' extensive analysis shows how Benzer sometimes condensed events in his relatively sketchy autobiographical reconstruction.

Trained as a solid-state physicist, Benzer was well prepared for quantitative analyses and genetic abstraction. In two years of training at the California Institute of Technology and a year at the Pasteur Institute in Paris, Benzer progressively penetrated the informal international phage network. Initially, he was deeply influenced by key players of the phage group, including Max Delbrück, Salvador Luria, André Lwoff, François Jacob, Alfred Hershey and Sidney Brenner. However, he progressively altered his pathway in the light of



Fine work: Seymour Benzer created a map of the gene with a resolution of just a few nucleotides.

his own results, as well as those of others — notably an experiment by Hershey and Martha Chase demonstrating the hereditary role of phage DNA, and James Watson and Francis Crick's work to develop the double-helix model of DNA.

Holmes ends his book with extensive accounts of the main public presentations by Benzer of his mapping results. About Benzer's landmark publication (*Proc. Natl Acad. Sci. USA* **41**, 344–354; 1955), Holmes states: "Like most modern scientific papers, Benzer's 'Fine Structure of a Genetic Region in Bacteriophage' is a logical reconstruction of the experiments, observations, and arguments supporting his conclusions. It has little narrative structure and does not purport to follow the investigative pathway from which it came."

Holmes patiently reconstructs the investigative path that led Benzer to draw the first high-resolution genetic map (down to a few nucleotides). The map supports the contention

that hereditary units are ordered linearly on the chromosome, and that genetic mutations and crossing-over experiments could be traced back to the underlying linear arrangement of DNA. Holmes says that Benzer isolated, characterized and crossed about a thousand T4 phage mutants, most of them in the limited rII region. This was possible thanks to the fantastic resolving power of Benzer's experimental system, which he progressively improved by using different types of mutants (extended deletions) to speed the mapping of novel mutants.

Benzer's results reached a wide audience through his contributions to two major conferences: the Brookhaven Symposium on Biology in 1955 and the McCollum-Pratt Institute Symposium on the Chemical Basis of Heredity at Johns Hopkins University in 1956. On the latter occasion, Benzer proposed three new terms — *cistron*, *recon* and *muton* — to solve the ambiguities associated with the term 'gene', which is considered altogether as a unit of function, recombination (crossing-over) and mutation. Although these terms (perhaps with the exception of 'cistron') did not become really popular among the emerging molecular-biology community, these distinctions helped to clarify the relationships between the different gene definitions. For a broader historical analysis, see *The Concept of the Gene in Development and Evolution*, edited by Peter Beurton, Raphael Falk and Hans-Jörg Rheinberger (Cambridge University Press, 2000).

Although technical in places, *Reconceiving the Gene* should nevertheless be accessible to a wide scientifically literate audience as it further introduces the broader context of classical and phage genetics. Holmes initially planned to go beyond Benzer's fine rII mapping and also cover his contributions towards solving the genetic code. This plan was impeded by illness, however, and Holmes died on 27 March 2003. Hopefully, his captivating book will stimulate other historians of biology to complete his project. ■

Denis Thieffry is in the Faculté des Sciences de Luminy, Université de la Méditerranée, 13288 Marseille, France.

S. BENZER

Leonardo's vision

A series of exhibitions across Europe show how Leonardo da Vinci linked art and science.

Stefano Grillo

For Leonardo da Vinci, painting did not mean merely copying the appearance of nature. Rather, it involved understanding nature's laws and using them to create a figurative world that enhances our awareness of reality. Leonardo's visual understanding meant that his scientific studies often took the form of beautiful drawings. His art and science sprang from the same source, providing a unified vision of the world and communicating a sense of wonder that is, according to Plato, the source of all philosophy and original investigation. Leonardo's will to 'know' was undoubtedly one of his strongest impulses, and his relentless questioning of how things work extended to all branches of the natural sciences of his day.

The Universal Leonardo project — an extraordinary collaboration of art and science historians, co-curated by *Nature* columnist Martin Kemp — does justice to Leonardo's broad interests in a series of events across Europe. The exhibition of drawings and manuscript pages at London's Victoria and Albert Museum describes how Leonardo 'thought on paper', for example, and the display at the Alte Pinakothek in Munich explores the scientific analyses of one of Leonardo's early paintings, the *Madonna and Child*.

Details of the exhibitions can be found in an extensive (and clever) website at www.universalleonardo.org, which also makes Leonardo's thought processes accessible to a wide audience. Its 'explore' page graphically highlights the links between Leonardo's art and science by identifying several themes running through the whole of his work.

The theme 'The Body of Earth', for instance, deals with the view — widespread in ancient and medieval times — that Earth is a macrocosm living according to the same laws as the human body. Blood circulation was seen to correspond to the movement of water in rivers (the veins of Earth) towards the sea (its heart). Leonardo accepted this model until late in life, when he saw inconsistencies that forced him to abandon it. He realized, for example, that the sea — unlike the heart — only received water; it did not pump it. He had tried to postulate the existence of subterranean rivers (arteries)



Leonardo's *Madonna and Child* and his drawing of an ox heart (right) can be seen at the various Universal Leonardo events.



bringing water to the top of mountains, but understandably failed, and the idea of the evaporation of water did not fit in with the macrocosm analogy.

One-quarter of the London exhibition is built around the macrocosm/microcosm theme, and illustrates it with drawings on hydrodynamics, cartography and anatomy, including some of the organs of animals. The drawing of the heart of an ox combines supreme beauty in the rendering of the plasticity of the form with penetrating and accurate observation of details, particularly of the coronary blood vessels.

Leonardo's work on palaeontology should also be understood in this context. At first sight, his astonishing studies of fossils give the impression that he was far ahead of his time. He refuted the contemporary idea that fossils formed following Noah's flood, and dismissed the neoplatonic theory that they grew inside rocks. He laid out in his unpublished notes a number of principles that were not clearly redefined by palaeontologists until the twentieth

century, such as the idea that growth rings in a fossil could be used to age it. But his aim was to find evidence to support a pre-modern, macrocosmic model of Earth, in which the elements were in constant flux, and the levels of water and mountains were constantly changing. As Stephen Jay Gould has pointed out, if we hope to understand Leonardo's true greatness we should abandon the metaphor of a "spaceman from a more advanced universe" (often accompanied by a condescending vision of the past), and look at his achievements within the context of his own time.

The Universal Leonardo project also allows present-day science to investigate Leonardo's works. X-ray radiography and infrared photography, for instance, were used to probe below the surface of the *Madonna and Child*. The results confirm that Leonardo

often changed his mind and modified his compositions at different stages of their execution. They also highlight the importance of geometry in his conception of form. In addition, the Munich exhibition displays 20 drawings of the *Madonna and Child* by a young Leonardo and his contemporaries.

One of the claims that Sigmund Freud made in his 1910 book *Leonardo da Vinci and a Memory of his Childhood* was that Leonardo had an unusually strong relationship with his mother. There is no documentary evidence for this, but the bond between mother and child in Leonardo's *Madonna* paintings and drawings appears deeper and stronger than in those by other artists. I like to think of his anatomical drawings of the fetus nestled in the protecting warmth of the womb (see *Nature* **396**, 25; 1998) as the scientific counterpart of this theme.

The scope for seeing connections between art and science in Leonardo's work is nearly boundless. By pointing some of them out, the Universal Leonardo project invites the visitor to move back and forth between these two fields — something Leonardo himself seems to have done continually.

Stefano Grillo is at the University of Perpignan and PROMES/CNRS, Tecnosud, 66100 Perpignan, France.

ALTE PINAKOTHEK, MUNICH

ROYAL COLLECTION/HER MAJESTY QUEEN ELIZABETH II

NEWS & VIEWS

MICROBIOLOGY

Death of a chaperone

Cesare Montecucco and Maurizio Molinari

To help their growth and spread, bacteria rely on virulence factors, many of which are toxic. One such factor is highly potent, as it attacks a key protein that 'chaperones' other proteins through their synthesis.

Microorganisms produce an almost infinite array of virulence factors. These molecules are shaped by evolution to alter selected physiological functions of the animal or plant hosts in such a way as to promote the microorganism's growth and/or spread in the world. Bacterial toxins are virulence factors so vicious that, in some cases, they can even kill the host. On page 548 of this issue, Paton *et al.*¹ describe the structure and enzymatic activity of a toxin, termed SubAB, which kills cells by digesting the BiP/GRP78 protein. BiP is a master regulator of the endoplasmic reticulum (ER; Fig. 1) — the membrane network where much of the synthesis and folding of proteins occurs in the cell — and SubAB is the first toxin known to hit this vital cellular target.

Paton and colleagues previously identified² this toxin in culture filtrates of a highly virulent strain of *Escherichia coli* that was responsible for the 1998 outbreak of haemolytic uraemic syndrome in southern Australia. This is a dangerous disease characterized by acute kidney failure, death of red blood cells (haemolytic anaemia) and diffuse aggregations of blood platelets that block the capillaries. There can also be neurological symptoms such as loss of coordination (ataxia) and peripheral paralysis. The bacteria that cause this disease are related to toxic strains of *E. coli* that cause gastrointestinal illnesses, and together these microbes are a major threat to human health³.

These *E. coli* strains have long been known to produce other toxins, the Shiga and Shiga-like toxins⁴, which consist of an enzymatic 'A' subunit and a group of five 'B' subunits, each capable of binding a glycolipid molecule in the cell-surface membrane. The Shiga A subunit is delivered to its site of action inside the cell using a membrane-transport process, whereby the patch of membrane to which the Shiga B subunits have bound is sucked inside the cell (endocytosis) and eventually joins up with the membranes of the ER. The A subunit is then shunted from inside the ER (the lumen) to the cytosol, where it damages an RNA that is essential for the cell's protein-synthesis machinery, thereby blocking protein synthesis and killing the cell⁴.

Paton *et al.*¹ find that SubAB also consists

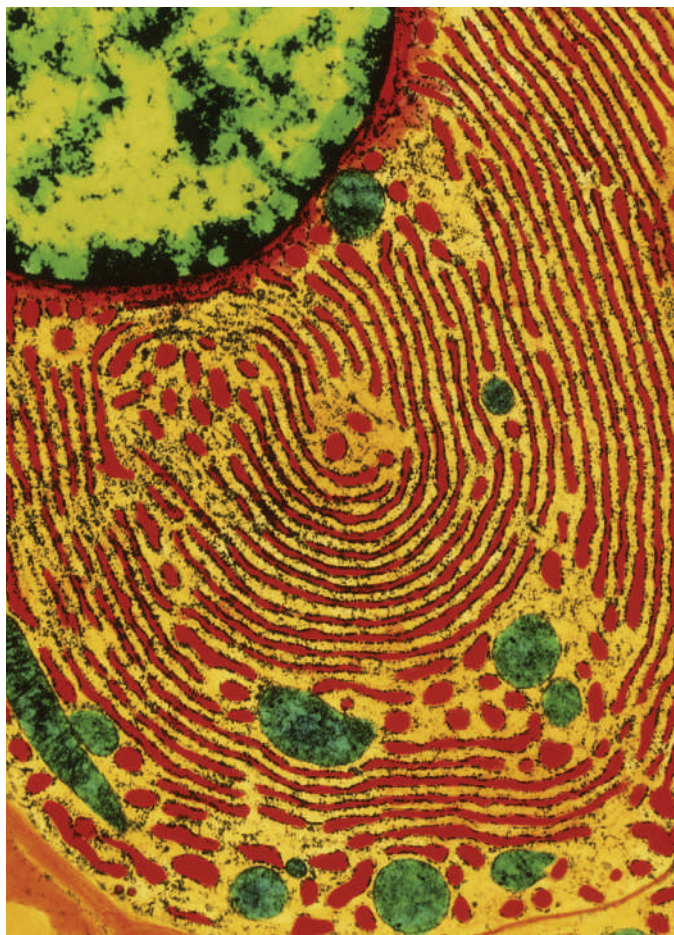


Figure 1 | The endoplasmic reticulum. This network of membranes is where much of the protein synthesis in the cell occurs (shown here at magnification approximately $\times 4,000$). Paton *et al.*¹ describe the action of a bacterial toxin called SubAB, which selectively attacks the BiP protein that is vital to the correct functioning of the endoplasmic reticulum. The cell nucleus is light green, mitochondria are dark green and the tiny black dots on the membranes are ribosomes, the molecular complexes that manufacture proteins.

CNR/SPL

of a binding pentamer of B units and an active A subunit, but, unlike the Shiga toxins, the A unit is very similar to a protein-digesting enzyme called subtilysin; it is these features that lead to the term SubAB. It took some proteomics, molecular biology and biochemistry for the authors to arrive at the remarkable description¹ of the structure and function of SubAB. This toxin binds to a type of cell-membrane glycolipid called ganglioside GM2, or similar ones, and then follows a transport pathway similar to that of the Shiga toxins to reach the ER lumen. Instead of being delivered into the cytosol, SubA remains in the lumen and attacks BiP, cutting the bond between two amino acids (Leu 416 and Leu 417) in an

exposed loop that links two BiP domains: the ATPase and the substrate-binding domains.

The toxin active site is characterized by an unusually deep cleft, and the authors suggest that this feature confers the precision required to find the exact cleavage site in BiP. However, it is likely that other SubAB regions are involved in the specific recognition of BiP, as is the case for the recognition of substrates by other bacterial toxins⁵.

BiP is a multifunctional protein^{6,7}. It maintains the permeability barrier of the ER lumen by sealing up inactive 'translocons' — protein-conducting channels in the ER. BiP aids translocation of growing protein chains into the ER lumen. It participates in folding

proteins and pairing them up with their partners, and it inhibits the aggregation of misfolded proteins. It contributes to the balance of calcium-ion concentrations in the ER. It helps to prepare misfolded proteins for destruction. And it activates an adaptive signalling pathway termed the 'unfolded protein response' that is fundamental to the well-being and development of cells, organs and tissues^{6,7}.

Notwithstanding all these functions, it is still rather surprising that the cleavage of BiP is so rapidly fatal to cells. Indeed, SubAB is at least ten times more toxic than the Shiga toxins are to cultured Vero cells. This is strong evidence for the central role of BiP in cell life. Also, it is possible that the toxin localization in the narrow volume inside the ER lumen increases the toxic effect, and it may be that there are even cells with a limited ER volume in which one molecule of toxin would be sufficient to cleave all BiP molecules present.

This cellular toxicity of SubAB fits with the characteristics of the diseases caused by the bacteria it comes from, as one would predict that cells very rich in gangliosides (such as neurons) or those that require a lot of protein synthesis (such as kidney epithelial cells and liver cells) would be worst affected by such a toxin. At the same time, the pathogenesis of haemolytic uraemic syndrome and the SubAB mechanism of action suggest that BiP must be particularly important in endothelial cells lining the blood vessels, because a major event in the disease is formation of blood clots in the capillaries, indicating damage to these vessels.

Another cell type that could be a particular target of SubAB is the plasma cell, which makes immunoglobulin proteins, the folding of which is specifically assisted and promoted by BiP. We suspect that these cells may be particularly sensitive to SubAB. If so, this could be highly relevant to the progress of many diseases, as the secretion of immunoglobulins by plasma cells is a major immune-defence mechanism, which would be impaired to the advantage of the bacterium. Paton *et al.*¹ also identify *E. coli* strains expressing both the Shiga and the SubAB toxins, raising the possibility that these two toxins cooperate in the final pathological action similarly to the anthrax oedema toxin and lethal factors in the pathogenesis of anthrax⁸.

SubAB will be a powerful tool for cell biologists to investigate the involvement of BiP in various cell functions. This is because of its exquisite specificity and the high level of evolutionary conservation of BiP — there is negligible sequence variation in the vicinity of the BiP cleavage site among proteins from multicellular animals. However, the hope is that, as with other bacterial toxins, SubAB may find some medical application. BiP at the surface of human prostate cancer cells enhances the spread of the cancer to other regions of the body, suggesting that this protein could be a cell-surface target for therapeutic use of the A subunit. Moreover, induction of BiP expression can be an unpleasant side effect of cancer therapies designed to

halt the formation of blood vessels, and it correlates with tumour aggressiveness, drug resistance, tumour recurrence and poor survival⁹. In such cases, SubAB may find a use in reducing such side effects during chemotherapy. ■

Cesare Montecucco is in the Department of Biomedical Sciences, University of Padua, 2-35121 Padua, Italy.

e-mail: cesare.montecucco@unipd.it

Maurizio Molinari is at the Institute for Research in Biomedicine, Via Vincenzo Vela 6, CH-6500 Bellinzona, Switzerland.

e-mail: maurizio.molinari@irb.unisi.ch

1. Paton, A. W. *et al.* *Nature* **443**, 548–552 (2006).
2. Paton, A. W., Srimanote, P., Talbot, U. M., Wang, H. & Paton, J. C. *J. Exp. Med.* **200**, 35–46 (2004).
3. Tarr, P. I., Gordon, C. A. & Chandler, W. L. *Lancet* **365**, 1073–1086 (2005).
4. Sandvig, K. *Toxicol.* **39**, 1629–1635 (2001).
5. Rossetto, O. *et al.* *Toxicol.* **39**, 27–41 (2001).
6. Hendershot, L. M. *Mt Sinai J. Med.* **71**, 289–297 (2004).
7. Luo, S., Mao, C., Lee, B. & Lee, A. S. *Mol. Cell. Biol.* **26**, 5688–5697 (2006).
8. Baldari, T., Tonello, F., Paccani, S. & Montecucco, C. *Trends Immunol.* **27**, 434–440 (2006).
9. Dong, D. *et al.* *Cancer Res.* **65**, 5785–5791 (2005).

ATOMIC PHYSICS

Quantum leap from light to atoms

Mikhail Lukin and Matthew Eisaman

Quantum-information networks use matter for long-term storage and light for long-distance transmission. Teleporting a quantum state from light onto matter has now been achieved.

A decade ago, the idea of teleportation was mere science fiction. But following the pioneering proposal of Bennett *et al.*¹, quantum teleportation has emerged as an important tool in quantum-information science. Quantum teleportation involves the disembodied transfer of a system's most basic feature — its quantum state — from one location to another. The first experimental demonstration of the phenomenon was the transfer of a quantum state of light onto another beam of light². More recently, the teleportation of a state between two single, trapped ions was demonstrated^{3,4}. On page 557 of this issue, Sherson *et al.*⁵ describe the transfer of a quantum state from a light pulse onto a collection of atoms, thereby achieving teleportation between two objects of different nature for the first time.

Quantum states are elusive objects, as they are easily destroyed by measurements. When only one copy of a state exists, only one measurement can be made. This single measurement cannot, even in principle, reveal complete information about the state. For this reason, quantum states cannot be copied, and one cannot transmit a quantum state by performing a direct measurement and then using a classical communication channel.

Quantum teleportation circumvents these problems by using correlations between physically separated quantum states — the phenomenon known as quantum entanglement — to transmit quantum states between distant locations. To teleport a quantum state between a sender (typically called Alice) and a receiver (typically called Bob), each party must possess half of a pair of entangled states. First, Alice performs a joint measurement, called a Bell measurement, on both the quantum state she intends to send and her half of the entangled pair. She sends the result to Bob over a classical

transmission line (a telephone will do). Bob can use his half of the entangled pair to reconstruct an exact copy of the initial state from the classically transmitted information.

The interest in quantum teleportation is not purely academic. The technique is a crucial element in so-called quantum repeaters, systems that can be used for long-distance quantum communication and secure quantum cryptography. Teleportation is especially helpful when direct communication between two remote locations is not possible because of losses in the connecting channel. Recently, the phenomenon has been shown to have an important role in fault-tolerant quantum computation, an error-correction mechanism for a noisy quantum computer⁶. Quantum teleportation is, in fact, likely to be an indispensable part of any robust quantum-information system.

Sherson and colleagues' advance⁵ is part of a broad effort now under way to use large ensembles of atoms as stable memory nodes for quantum communication^{7,8}. In this approach, long-lived spin states of alkali atoms are used to store quantum states, while photons are used for communication. The memory nodes are connected to the communication channels through controlled light-matter interactions, using established techniques from the field of high-resolution laser spectroscopy. As light pulses typically interact with many atoms at once, the photonic quantum states are stored in collective spin excitations of atomic ensembles. Recent experimental advances within this framework include the entanglement of two remote atomic ensembles⁹, and the generation, storage and retrieval of single-photon pulses in primitive, two-node quantum networks^{10,11}.

The present paper⁵ is based on several conceptual and technical advances developed over the past few years in a close collaboration

between experimentalists led by Eugene Polzik and theorists led by Ignacio Cirac. The authors begin by creating entanglement between a light beam and approximately 10^{12} caesium atoms by passing a strong laser beam through an ensemble of spin-polarized atoms in a magnetic field. To establish quantum entanglement, a quantum-mechanical version of the Faraday effect is used, in which the polarization of light is rotated as it passes through a medium. In the experiment, this effect causes the fluctuations in polarization of the transmitted light to become correlated with the spin fluctuations in the caesium ensemble.

The authors then use a remote beamsplitter to mix the transmitted beam with a weak laser pulse whose quantum state is to be teleported. A Bell-type measurement of the so-called canonical variables (roughly corresponding to the amplitudes and phases) of the mixed pulses follows. The results of the measurement are transmitted back to the atomic ensemble through a classical feedback loop, where they are used to apply a magnetic field to the atomic ensemble. As the ensemble already 'knows' the state of the transmitted light with which the weak pulse was mixed, the effect is to map the initial quantum state of the weak laser pulse onto the collective spin excitation of the ensemble.

To verify that the teleportation was successful, the authors repeat the experiment many times to gather statistical information about the differences between the initial state of the weak light pulse and the final state of the atoms. The resulting figure of merit, called the fidelity, indicates that the observed teleportation is not perfect; but it is higher than would be possible with any classical approach that does not use entanglement. Moreover, the teleportation occurs 'on demand', at a time determined by the scientists.

Although the present experiment is an

important proof-of-principle demonstration, much work remains to determine whether such techniques can become practical tools. For example, it is likely that in useful quantum-information systems, photonic quantum bits, or qubits (for example, the two polarization states, horizontal and vertical, of a single photon) will be required, rather than the weak, coherent states used by Sherson *et al.*⁵. The teleportation of single-photon qubits does seem feasible with the present techniques, but reaching high fidelities will be difficult. Likewise, the powerful applications of teleportation in fault-tolerant systems require local operations to be performed with high efficiency and high fidelity, something that is often difficult in the absence of the large, nonlinear responses found in ensemble-based systems. Thus, further conceptual advances in designing efficient systems with limited resources are needed.

Beyond potential applications, the experiment of Sherson *et al.*⁵ demonstrates an exceptional degree of quantum control over light and matter. Such control is much needed as the field of experimental quantum-information science matures. ■

Mikhail Lukin and Matthew Eisaman are in the Department of Physics, Harvard University, 17 Oxford Street, Cambridge, Massachusetts 02138, USA.
e-mail: lukin@physics.harvard.edu

1. Bennett, C. H. *et al.* *Phys. Rev. Lett.* **70**, 1895–1899 (1993).
2. Bouwmeester, D. *et al.* *Nature* **390**, 575–579 (1997).
3. Barrett, M. D. *et al.* *Nature* **429**, 737–739 (2004).
4. Riebe, M. *et al.* *Nature* **429**, 734–737 (2004).
5. Sherson, J. F. *et al.* *Nature* **443**, 557–560 (2006).
6. Knill, E. *Nature* **434**, 39–44 (2005).
7. Lukin, M. D. *Rev. Mod. Phys.* **75**, 457–472 (2003).
8. Sherson, J., Julsgaard, B. & Polzik, E. S. *Adv. Atom. Mol. Opt. Phys.* (in the press); preprint available at www.arxiv.org/quant-ph/0601186 (2006).
9. Chou, C. W. *et al.* *Nature* **438**, 828–832 (2005).
10. Chanelière, T. *et al.* *Nature* **438**, 833–836 (2005).
11. Eisaman, M. D. *et al.* *Nature* **438**, 837–841 (2005).



50 YEARS AGO

"Marking of tsetse flies for their detection at night" — Most testse field-work so far has been based on day-time observations of active flies. If work is to be done on resting flies, it will be advantageous to have an easy method of locating them. It is known that mosquitoes can be dusted with non-toxic materials which fluoresce in ultra-violet light at night, and using this method, mosquitoes can be watched at a distance of about 10 ft. The insects are not sensitive to ultra-violet radiations...

'Dayglo' (silica-based organic dyes in ethyl cellulose) powders mixed with a binder and solvent produced a satisfactory paint and were used in a series of tests on house and tsetse flies... Paint was applied with a sharpened match-stick on the centre of the thorax, covering it without fouling head or wings. At 15 ft., treated flies shone brilliantly in the ultra-violet beam and the maximum range was 20 ft. **G. R. Jewell**
From *Nature* 6 October 1956.

100 YEARS AGO

"The genesis of the inventor" (*Erfindung und Erfinder* by A. du Bois-Reymond) — The author's analysis of invention and inventors leads to the conclusion that neither need, nor chance, nor the lack of necessities in surrounding life suffices to draw out the inventor. Instead of solving the problem by philosophic deduction from generalities, he descends to the particulars of the Patent Office, and concludes that inventors can be subdivided into three classes:— first, the intuitive genius, or, as Herbert Spencer would have said, the man who can do with little trouble that which cannot be done by the ordinary man with any amount of trouble; secondly, the technical man, well acquainted with his work, who follows in the wake of the intuitive genius, and is largely inspired by him; thirdly, the layman, whose special province seems to be feeding-bottles. From *Nature* 4 October 1906.

CELL CYCLE

Complex evolution

Gavin Sherlock

Cell division is fundamental to life, and so might be expected to have changed little during evolution. Data from four species show that the genes involved can vary, but the regulation of complexes is a common theme.

There are many core biological processes for which different organisms use essentially the same proteins to carry out certain tasks. These equivalent proteins, known as orthologues, are mostly very similar in terms of their protein sequence. However, what remains much less clear is whether these orthologues are regulated in a similar manner in different organisms — either at the level of transcription from the gene, or at the level of modification and degradation of their gene products. On

page 594 of this issue, Jensen *et al.*¹ use genome-wide data from four different organisms to investigate the periodic gene transcription that is associated with the cell-division cycle, and to determine how individual members of crucial protein complexes are regulated.

Using the publicly available data sets for budding yeast (*Saccharomyces cerevisiae*), fission yeast (*Schizosaccharomyces pombe*), thale cress (*Arabidopsis thaliana*) and human cells, the authors show that most of the genes

50 & 100 YEARS AGO

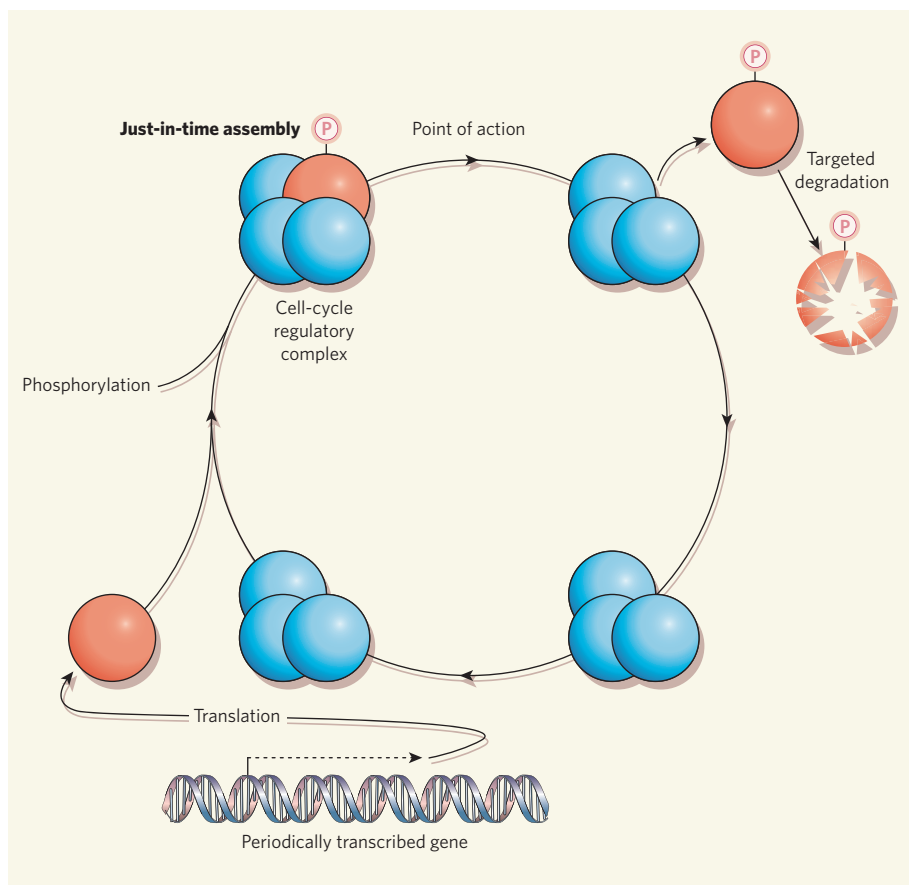


Figure 1 | 'Just-in-time' assembly of cell-cycle complexes. According to this idea, the complex is fully assembled only just before it is required to act. The subunit encoded by a gene that is transcribed periodically is available briefly, and is activated by phosphorylation, then targeted for degradation immediately after the complex has performed its task. Jensen *et al.*¹ show that, in support of this, genes that are transcribed periodically during the cell cycle produce proteins that are more likely to be phosphorylated and/or selectively degraded than other members of cell-cycle complexes.

expressed periodically do not have orthologues in all four organisms. And for the orthologues that are conserved across all four, the timing of their expression during the cell cycle is generally not conserved. Given the previously established conservation of many components of the core cell-cycle machinery, and even the ability of human proteins to stand in for mutated cell-cycle proteins in yeast², these widespread differences in both the identity and timing of transcription of the periodic genes seem unexpected.

Jensen *et al.* showed previously³ that many protein complexes that function in the yeast cell cycle have some members that are transcribed periodically and some that are transcribed constantly (constitutively). Their current work shows that, although this observation remains generally true across organisms, the identity of the periodically transcribed and constitutive members of a given complex varies between different organisms. Furthermore, those proteins whose transcripts are periodic are more likely to undergo some form of post-translational regulation than are others in the complex: either they acquire a phosphate group — a common means of regulating or activating proteins — and/or they are targeted for

degradation. This suggests that such a protein is active and/or available only for a short time, and that it joins the constitutively transcribed proteins, which are available continuously, just briefly while the complex carries out its function (Fig. 1). This implies that such 'just-in-time' assembly³ is more important for the correct functioning of the complexes than is the actual identity of the proteins whose transcription is periodic. In this scenario, the periodic protein acts as a trigger for the functioning of the complex, and it matters less what the periodic protein is than that it is present at the right time to fire up the complex.

Thus, post-translational control of proteins encoded by genes that are periodically regulated seems to be crucial for maximal control of the cell-cycle complexes. Regulation of the periodic transcription of these genes must therefore have coevolved with their post-translational modifications. It is not clear how fast transcriptional and post-translational regulation can coevolve, nor what the situation in the ancestral common organism was. However, it is known that transcriptional networks can be readily remodelled during *in vitro* evolution studies within only a few hundred generations under specific selective pressures⁴.

A key question is how such transcriptional diversity between organisms arises, and how phosphorylation and degradation might coevolve with transcriptional changes. Data are accumulating showing that there is considerable transcriptional diversity even within different individuals of the same species. For instance, in the progeny generated from a reproductive cross between different strains of *S. cerevisiae*, thousands of genes show evidence of heritable variation⁵, suggesting that there is great transcriptional diversity on which evolution can act. Some of this diversity is due to changes in a single nucleotide base (polymorphisms) lying near or in the coding sequence of a particular gene, when such changes influence that gene's expression (*cis* regulation). But much of the heritable variation seems to be due to polymorphisms in genes that affect the transcription of other genes (*trans* regulation), including many that encode proteins involved in modifying the way genes are packaged as chromatin⁶. Nonetheless, much of the variation remains unexplained, even by the best models⁶.

Recent gene-expression analyses comparing *S. cerevisiae* and the yeast *Candida albicans* showed that, for many groups of genes, certain genes are expressed together in one of the organisms but not in the other^{7,8}. Indeed, it was noted specifically that a group of genes involved in the transition from one stage of the cell cycle to the next (S phase to mitosis) are tightly co-regulated in *C. albicans*, but not in *S. cerevisiae*⁷. Differences in the transcriptional responses to stress between closely related *Saccharomyces* species have also been found⁹. Clearly, there is great diversity in transcriptional networks and their regulation, both within and between species — what we don't yet understand is how post-translational regulation may coevolve with such transcriptional changes. Currently, the data are too limited, and the sampled taxa far too sparse, for us to understand these processes fully. Additional data sets that assay various biological processes, sampling diverse taxa, as well as densely sampling particular taxa and species, will undoubtedly yield answers to these questions, and disentangle this complex evolution. ■

Gavin Sherlock is in the Department of Genetics, 300 Pasteur Drive, Stanford University Medical School, Stanford, California 94305-5120, USA. e-mail: sherlock@genome.stanford.edu

- Jensen, L. J., Jensen, T. S., de Lichtenberg, U., Brunak, S. & Bork, P. *Nature* **443**, 594–597 (2006).
- Lee, M. G. & Nurse, P. *Nature* **327**, 31–35 (1987).
- de Lichtenberg, U., Jensen, L. J., Brunak, S. & Bork, P. *Science* **307**, 724–727 (2005).
- Ferea, T. L., Botstein, D., Brown, P. O. & Rosenzweig, R. F. *Proc. Natl Acad. Sci. USA* **96**, 9721–9726 (1999).
- Brem, R. B. & Kruglyak, L. *Proc. Natl Acad. Sci. USA* **102**, 1572–1577 (2005).
- Lee, S. I., Pe'er, D., Dudley, A. M., Church, G. M. & Koller, D. *Proc. Natl Acad. Sci. USA* **103**, 14062–14067 (2006).
- Ihmels, J., Bergmann, S., Berman, J. & Barkai, N. *PLoS Genet.* **1**, e39 (2005).
- Ihmels, J. *et al. Science* **309**, 938–940 (2005).
- Tirosh, I., Weinberger, A., Carmi, M. & Barkai, N. *Nature Genet.* **38**, 830–834 (2006).

EARTH SCIENCE

Lost lithium found

Elisabeth Widom

Lithium isotopes provide a fingerprint of recycled material in Earth's upper mantle. But this fingerprint is different from what had been expected. So do we need to reassess our ideas about how the upper mantle evolves?

Earth's interior is thought to evolve through two complementary processes: the production of crust at the surface by convection and resultant partial melting of the underlying mantle, and the return of the crust to the mantle through subduction. The precise nature of this recycling process, including the ultimate fate of recycled crust and its influence on the evolution of Earth's upper mantle, is unclear. On page 565 of this issue, Elliott *et al.*¹ present lithium isotope results for rocks from the East Pacific Rise, a region of oceanic crust production that runs parallel to the west coast of the Americas. The data provide strong evidence for crustal recycling, but challenge the notion of simple re-enrichment and mixing of oceanic crust into the upper mantle.

The shallow mantle that feeds mid-ocean ridges such as the East Pacific Rise is, typically, chemically depleted; that is, it has lost much of its original complement of 'incompatible' elements, which have been scavenged by mantle melts during previous episodes of melting and crust formation². Rocks known as mid-ocean-ridge basalts (MORBs) record this depleted character. However, basalts from some regions of the shallow mantle reservoir show compositional and isotopic anomalies indicating that previously depleted mantle has been re-enriched. This re-enrichment has frequently been attributed to oceanic crust that, after ageing and alteration at Earth's surface, has been subducted and mixed back into the mantle³.

Subduction occurs where a region of the oceanic lithosphere (a rigid layer, or 'plate' that includes the basaltic oceanic crust and uppermost mantle) descends into the mantle beneath an overriding lithospheric plate. During subduction, the lithosphere becomes 'dewatered' as its hydrous fluids are released and permeate the overlying wedge of mantle. This process causes the mantle to melt and produces characteristic arc volcanism at the surface (Fig. 1). Compelling evidence for this recycling of oceanic crust comes from what are essentially computer-assisted tomography (CAT) scans of Earth's interior. These have revealed many instances of subducted lithospheric slabs in the upper mantle that, in some cases, extend all the way to the boundary with Earth's core⁴; the ultimate fate of these slabs is unknown.

Although the evidence for slab recycling seems conclusive, it is less clear that recycled oceanic crust mixes into and so re-enriches upper mantle that has become depleted in

incompatible elements by earlier episodes of melting. Some MORBs, and hence their mantle sources, are relatively enriched in incompatible elements such as potassium, thorium, uranium and light rare-earth elements. These enrichments could, in theory, result from mantle regions that have undergone mixing with recycled oceanic crust⁵. But the extent to which incompatible elements are lost from a subducted slab during dewatering remains uncertain⁶. Predicting the precise chemical fingerprint of a mixture of depleted mantle and recycled oceanic crust is therefore difficult. Alternative enrichment mechanisms, such as replenishment by melts derived from deeper in the mantle, cannot readily be excluded.

Isotopes of light elements have unique potential as diagnostic tracers of crustal recycling processes. Such isotopes, for example the lithium isotopes ⁶Li and ⁷Li, behave similarly to one another at high temperatures. Owing to their large relative-mass difference, however, they can behave differently during low-temperature processes at or near Earth's

surface, and become separated (fractionated) from one another.

The ratio ⁷Li/⁶Li does not vary much in the depleted mantle, or during mantle melting and the subsequent crystallization of basaltic magma⁷. By contrast, extreme variations in the lithium isotope ratio — both above and below normal mantle values — occur in the oceanic lithosphere, as it reacts extensively with isotopically heavy sea water (enriched in ⁷Li relative to ⁶Li) during and after its formation at mid-ocean ridges. Further fractionation of lithium isotopes occurs in the early stages of subduction, because ⁷Li has a greater affinity for the fluid phase during dewatering. Thus, subducted slabs should be isotopically lighter than depleted mantle⁸. Regardless of whether or not this is true, variation in lithium isotopes in basaltic magmas requires the presence of recycled crustal material, or of its derivatives, in enriched mantle regions (assuming that we can rule out interactions with crust near the surface during the ascent of magma).

Elliott *et al.*¹ use isotope ratio measurements of particularly high precision to investigate the origin of the enriched mantle that generates enriched MORBs at the East Pacific Rise. The authors' results reveal for the first time correlations between lithium isotopes and other chemical and isotopic indices of mantle enrichment in MORBs. These correlations provide strong evidence for the involvement of crustal recycling processes in the formation of enriched domains in the upper mantle. But the enriched MORB samples have isotopically

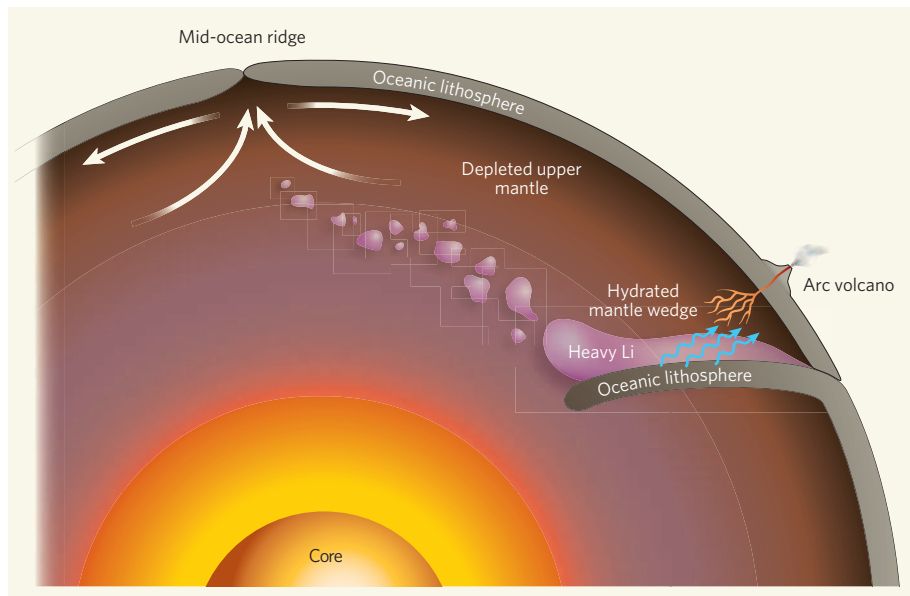


Figure 1 | The view from within. During various episodes of melting and crust formation, the upper mantle generally becomes depleted in so-called incompatible elements. But the mantle sources of some mid-ocean-ridge basalts (MORBs) are unusually enriched in such elements, and it had been suggested that mixing with subducted oceanic crust was the cause. Elliott *et al.*¹ suggest an alternative mechanism to account for the isotopically heavy lithium composition of enriched MORBs from the East Pacific Rise. At a subduction zone, one plate of oceanic lithosphere dives under another plate. This is associated with the release of fluid ('dewatering'; blue arrows) into the overlying mantle wedge, as well as melting that produces arc volcanism at the surface. In the authors' model, it is part of the hydrated mantle wedge above the subducting lithosphere, which is initially viscously coupled to the downgoing slab, that eventually frees itself and mixes into surrounding depleted mantle. This produces enriched mantle domains that can be sampled during melting beneath the mid-ocean ridges.

NANOTECHNOLOGY

Downsizing SQUIDS

They sound fishy, but SQUIDs — superconducting quantum interference devices — are ultra-sensitive gadgets used for measuring the strength of magnetic fields. These 'magnetometers' have widespread applications ranging from materials science to medicine. Writing in *Nature Nanotechnology*, Cleuziou, Wernsdorfer and colleagues describe a minuscule version of this useful device (J.-P. Cleuziou *et al.* *Nature Nanotech.* **1**, 53–59; 2006).

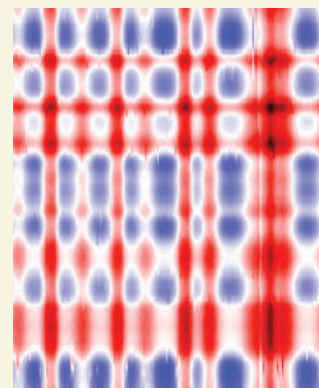
The 'nano-SQUID' is made of a square loop of 50-nm-thick aluminium wire, which becomes

superconducting at temperatures below 1.2 K. Insulating junctions disrupt the loop in two places, with each junction consisting of a single-walled carbon nanotube. The maximum supercurrent flowing through the loop varies with changes in the magnetic flux (the magnetic field multiplied by the area) passing through the loop. This response is what makes SQUIDs such sensitive magnetometers.

The nanotube junctions behave like 'quantum dots', which have discrete electronic energy levels. The alignment of these energy levels with those of the superconducting

wire can be controlled by applying a voltage near to either — or both — of the nanotube junctions. In this way, the voltage effectively turns the current on (energy levels aligned) and off (misaligned). The two-dimensional colour plot shown on the right indicates the peaks (red) and valleys (blue) in the conductance of the device as the junctions are tuned through their on and off states.

So why make a SQUID that has carbon-nanotube junctions? The ultimate goal is to detect changes in direction of the magnetic moment of a single molecule or nanoparticle. This could be achieved by placing the nanoparticle directly on one of the nanotubes, optimizing the magnetic



flux passing through the loop and so maximizing the device's sensitivity. Although such a measurement has not yet been made, it is well within the sensitivity of this tiniest of SQUIDs.

Jessica Thomas

heavier lithium than has normal depleted mantle. This is surprising, as they would have been expected to be isotopically lighter if the enriched mantle were formed by simple mixing of subducted oceanic crust whose ^7Li had largely escaped upwards into the mantle wedge during dewatering.

So where did this heavy lithium come from? Elliott *et al.*¹ propose that, in fact, it is the same heavy lithium that is released during dewatering of a subducting slab. In their model, the mantle wedge immediately above a subducting slab is initially relatively cold and viscously coupled to the downgoing slab. As it warms up, it eventually peels away from the denser basaltic slab, and remixes with surrounding depleted mantle to produce enriched domains (Fig. 1). Although the ultimate fate of the slab remains unknown, this model provides a solution for another mystery: the surprising absence of heavy lithium isotope signatures in most arc volcanoes, despite abundant evidence that their mantle source regions contain fluids from the subducting slab⁹. Elliott and colleagues' model supports previous speculation that lithium is rapidly scavenged from slab fluid as it percolates through the mantle wedge, owing to an affinity of lithium for magnesium silicate minerals in the mantle⁹.

Confirmation of this model will require further study of how lithium behaves during dewatering of subducting oceanic lithosphere. Limited work on eclogites — rocks thought to be fragments of previously subducted oceanic crust — show a wide range of lithium isotope signatures, but generally support the notion that isotopically heavy lithium is released during slab dewatering⁸. Although it remains to be shown that subducted oceanic crust is uniformly light in its lithium isotopic signature, Elliott and colleagues' demonstration of lithium isotopic variations in enriched MORBs remains strong evidence for the importance of

crustal recycling — in one form or another — in the evolution of Earth's upper mantle. ■

Elisabeth Widom is in the Department of Geology, Miami University, Oxford, Ohio 45056, USA.

e-mail: widome@muohio.edu

1. Elliott, T., Thomas, A., Jeffcoate, A. & Niu, Y. *Nature* **443**, 565–568 (2006).
2. Hofmann, A. W. *Nature* **385**, 219–229 (1997).

3. Allègre, C. J. & Turcotte, D. L. *Nature* **323**, 123–127 (1986).
4. Hutko, A. R., Lay, T., Garner, E. J. & Revenaugh, J. *Nature* **441**, 333–336 (2006).
5. Hanan, B. & Graham, D. *Science* **272**, 991–994 (1996).
6. Tatsumi, Y. & Kogiso, T. *Earth Planet. Sci. Lett.* **148**, 207–221 (1997).
7. Tomascak, P. B. *Rev. Mineral. Geochem.* **55**, 153–195 (2004).
8. Zack, T., Tomascak, P. B., Rudnick, R. L., Dalpe, C. & McDonough, W. F. *Earth Planet. Sci. Lett.* **208**, 279–290 (2003).
9. Tomascak, P. B., Widom, E., Benton, L. D., Goldstein, S. L. & Ryan, J. G. *Earth Planet. Sci. Lett.* **196**, 227–238 (2002).

MICROBIOLOGY

Resurrecting a broken genome

Susan T. Lovett

A remarkable bacterium can survive extraordinary doses of ionizing radiation that shatter its genome into thousands of pieces. How does it accurately reassemble these DNA fragments into an intact genome?

Aficionados of DNA repair can't help but be impressed by the small but mighty bacterium *Deinococcus radiodurans*. When its genome is smashed into thousands of little pieces by ionizing radiation, *D. radiodurans* can reconstruct it without error¹. This lowly organism, which thrives in nuclear waste², can certainly teach us a thing or two about DNA repair. On page 569 of this issue, Zahradka *et al.*³ reveal how this bacterium manages to rebuild its shattered chromosomes*.

Deinococcus radiodurans was isolated in 1956 as a contaminant in canned meat that had been treated with 4,000 gray of γ -rays⁴, a dose many hundreds of times greater than that required to kill 'normal' bacteria such as *Escherichia coli*. Given that no terrestrial environment

*This article and the paper concerned³ were published online on 27 September 2006.

experiences such enormous doses of radiation, how on Earth has this radiation resistance evolved? The answer is simple: radiation resistance correlates with resistance to desiccation⁵. *Deinococci* are aerobic bacteria found in soil around the globe. Prolonged dehydration, like ionizing radiation, induces potentially lethal double-strand DNA breaks in the bacterial genome. Presumably, *D. radiodurans* has evolved to survive dry periods in its soil environment by developing efficient mechanisms for repairing chromosome breaks. Extremely radiation-resistant bacteria from surprisingly diverse lineages are readily isolated from desert soil^{6,7}, so this trait seems to have arisen many times as a response to arid environments.

In the 50 years following the discovery of extreme radiation resistance in bacteria, many people have pondered over its mechanism

(or mechanisms). It is known that *D. radiodurans* is not impervious to chromosome breakage: the average dose of radiation needed to kill the bacterium does so by fragmenting its genome into pieces of 10 kilobases (kb). Compare this with *E. coli*, which dies when its genome is broken into 500-kb pieces⁸. On this basis, *D. radiodurans* does seem to have a superior ability to repair double-strand breaks in DNA.

Another puzzle is how *D. radiodurans* reconstructs its genome without confusion between repeated DNA sequences that occur in multiple copies in its chromosomes; such confusion would introduce lethal genetic rearrangements. The genome of *D. radiodurans* strain R1 consists of two circular chromosomes (one 2.65 megabases in size, the other 0.45 megabases) and two 'mini-chromosomes' known as plasmids (177 kb and 46 kb). Compared with *E. coli*, *D. radiodurans* has more, rather than fewer, repeated-sequence elements⁹, the longest of which is 1.3 kb and is present at 13 locations in the genome. Also, like many bacteria, a *D. radiodurans* cell produces several identical copies of its chromosomes during its cell-division cycle, and thus has at any one time between four and twelve redundant copies of the genome that can be used as templates to help construct an intact genome after damage.

Zahradka *et al.*⁵ now reveal the molecular mechanism of genome restoration in *D. radiodurans*. The repair of chromosomes that have been fragmented into 20–30-kb pieces is accompanied by extensive DNA synthesis. Unlike normal semiconservative DNA replication, in which only one strand of each resulting DNA duplex is synthesized anew, the *D. radiodurans* repair mechanism synthesizes both strands of new DNA duplexes in a process mediated by the enzyme DNA polymerase A (PolA). A second step then assembles an intact chromosome, assisted by the universal recombination protein RecA. The authors' analyses³ suggest that repaired chromosomes are a patchwork of old and new 20–30-kb blocks of DNA that have been joined together.

Transient accumulation of single-stranded DNA was detected during the repair process, and taken together, these results³ are consistent with a 'synthesis-dependent strand-annealing' mechanism (Fig. 1). This mechanism has been proposed to explain how DNA breaks made by P element transposons — DNA sequences that can 'jump' around to different positions in the genome — are repaired in fruitflies¹⁰. In this process, a broken DNA end is digested by a nuclease enzyme to leave a single-stranded tail. This tail then anneals to a strand of complementary DNA in an intact chromosome and acts as a primer for DNA synthesis, producing a long single strand of newly synthesized DNA. A similar reaction from a nearby break extends the synthesis in the opposite direction, generating a single DNA strand complementary to the first one. Subsequent annealing of these two single

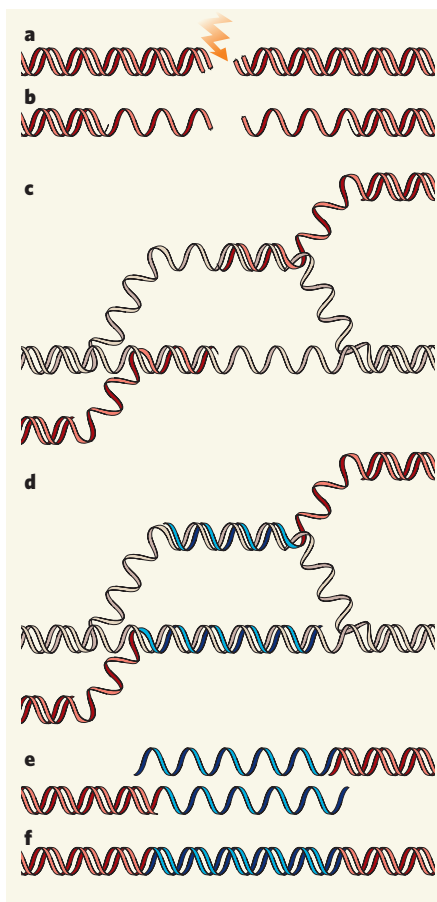


Figure 1 | DNA repair. The bacterium *Deinococcus radiodurans* has a remarkable ability to repair its genome following radiation damage, using a mechanism consistent with synthesis-dependent strand-annealing. **a**, Radiation or desiccation produces double-strand DNA breaks. **b**, Nuclease enzymes partly digest the broken DNA to expose single-stranded ends. **c**, The single strands then anneal to complementary DNA in an intact chromosome. **d**, Using the chromosome as a template, the invading single strands prime DNA synthesis (blue lines). **e**, The extended single strands unwind from the duplex DNA, producing long single-stranded tails of newly synthesized DNA. **f**, The single strands anneal together, so that fragments from the original damaged genome are linked with freshly synthesized DNA.

strands generates a duplex of newly synthesized DNA that has joined two fragments together. The broken fragments become joined up in this way to form long linear copies of the chromosome. These are then reduced in a final step to an intact circular copy by a process known as homologous recombination.

So how does *D. radiodurans* ensure quality control during chromosome reconstruction? There are two stages that are vulnerable to errors — the initial annealing of a single DNA strand to an intact chromosome to initiate DNA synthesis, and the subsequent annealing of the newly synthesized DNA strands. If either of these pairing reactions uses a repeated DNA sequence element from the wrong part of the genome, lethal genetic rearrangements can result. Proteins

that unwind duplex DNA may abort pairing between short duplicated elements in *D. radiodurans*, as they do during DNA repair in *E. coli*. Or perhaps proteins that bind to repeated DNA sequences prevent them from annealing. The extraordinary survival of *D. radiodurans* following genome fragmentation may well result from its special ability to prevent errors during the synthesis-dependent reassembly of DNA.

Now that the basic mechanism of the repair reaction in *D. radiodurans* has been revealed, it will be crucial to understand how individual proteins promote repair efficiency and fidelity. Although some factors specific to *D. radiodurans* have been isolated that promote double-strand-break repair¹¹, the genome sequence of this bacterium reveals an unexpected complement of DNA-repair and recombination proteins typical of normally radiation-sensitive bacteria like *E. coli*. The nucleases RecBC and ExoI are, however, notable by their absence; this deficiency is probably critical for the preservation of DNA ends during the genome-assembly process.

After years of speculation about the various exotic ways in which *D. radiodurans* might facilitate DNA repair, this bacterium seems, somewhat surprisingly, to employ a mechanism believed to occur in all cells. The apparently independent and recurrent evolution of resistance to radiation and desiccation in bacteria raises the possibility that these unusual properties may not require any fundamentally distinct mechanism, just some tweaking of basic biological mechanisms already present. The success of *D. radiodurans*, like that of a well-trained athlete, may not result from a singular spectacular talent, but from its ability to do every little thing just right. ■

Susan T. Lovett is in the Department of Biology, Brandeis University, Waltham, Massachusetts 02454-9110, USA.

e-mail: lovett@brandeis.edu

1. Battista, J. R., Earl, A. M. & Park, M. J. *Trends Microbiol.* **7**, 362–365 (1999).
2. Lange, C. C., Wackett, L. P., Minton, K. W. & Daly, M. J. *Nature Biotechnol.* **16**, 929–933 (1998).
3. Zahradka, K. *et al.* *Nature* **443**, 569–573 (2006).
4. Anderson, A. W., Nordon, H. C., Cain, R. F., Parrish, G. & Duggan, D. *Food Technol.* **10**, 575–578 (1956).
5. Mattimore, V. & Battista, J. R. *J. Bacteriol.* **178**, 633–637 (1996).
6. Chanal, A. *et al.* *Environ. Microbiol.* **8**, 514–525 (2006).
7. Rainey, F. A. *et al.* *Appl. Environ. Microbiol.* **71**, 5225–5235 (2005).
8. Cox, M. M. & Battista, J. R. *Nature Rev. Microbiol.* **3**, 882–892 (2005).
9. Makarova, K. S. *et al.* *Microbiol. Mol. Biol. Rev.* **65**, 44–79 (2001).
10. Nassif, N., Penney, J., Pal, S., Engels, W. R. & Gloor, G. B. *Mol. Cell. Biol.* **14**, 1613–1625 (1994).
11. Tanaka, M. *et al.* *Genetics* **168**, 21–33 (2004).

Correction

In the News and Views article "Hydrogen at the flick of a switch" by Masanori Takimoto and Zhaomin Hou (*Nature* **443**, 400–401; 2006), the e-mail address for Zhaomin Hou was incorrect. The correct address is houz@riken.jp

NEWS & VIEWS FEATURE



Transposable elements affect maize colour.

GENETICS

Junk DNA as an evolutionary force

Christian Biémont and Cristina Vieira

Transposable elements were long dismissed as useless, but they are emerging as major players in evolution. Their interactions with the genome and the environment affect how genes are translated into physical traits.

Transposable elements (TEs) — commonly called ‘jumping genes’ — are stretches of DNA that move around the genome of a cell, and the genomes of many higher organisms are cluttered with numerous copies of these enigmatic elements. They were discovered by Barbara McClintock in the 1950s (Box 1), but it has taken half a century to begin to understand how they act and the effects they can have. It is emerging that these elements have had a significant influence on the evolution of genomes, particularly by controlling gene activity.

The elements contain in their sequence all the instructions needed to cut themselves out of their host DNA and splice themselves into another spot. But they are not always benign ‘junk’ DNA — they can insert into genes or gene regulatory elements, potentially disrupting the gene’s function, and they can trigger chromosome rearrangements. So, even though most copies are selectively neutral and not in themselves damaging, they have long been consid-

ered as predominantly harmful to their hosts, as they can contribute to the appearance of mutations, some of which can result in disease.

But TEs do not always have adverse effects, and their mutational activities contribute to the genetic diversity of the organism. Indeed, some TEs have been domesticated by their host genome, acting as genes or gene regulatory elements, and as a result constitute a source of genetic innovation for the organism^{1,2}. Progress in understanding how these elements are regulated is bringing an appreciation of how an individual’s environment can affect the expression of their genetic complement to produce their own particular characteristics (phenotypes), such as physical appearance, behaviour, susceptibility to disease and even neuronal function.

Diversity

Transposable elements are scattered throughout the genomes of many plants and animals, and can form a large proportion of the genome

size (Table 1). There are two main classes of TE: DNA transposons, which act through a DNA intermediate and multiply by using the host cell’s replication machinery, and retrotransposons, which act through an RNA intermediate. Retrotransposons are further subdivided into those that have ‘long terminal repeats’ at their ends (LTR retrotransposons) and those that do not (non-LTR retrotransposons; Box 2). In addition, TEs with composite structures are continually being discovered, illustrating the enormous flexibility of these elements. For example, DNA transposon-like elements called helitron rolling-circle elements were recently found to be responsible for copying various gene segments into new locations in the maize genome, generating a huge diversity among individual maize plants³.

When LTR retrotransposons are excised from the genome, they leave behind an LTR sequence. Some genomes, particularly those of plants, are full of these lone LTRs. Because they

PIXOLIO/LAMY

Box 1 | Transposable elements in gene regulation

Barbara McClintock discovered mobile genetic elements in the 1940s during her studies of maize, work for which she received the 1983 Nobel Prize in Physiology or Medicine. She observed that the patterns of colour in maize kernels changed in various breeding crosses, and she interpreted her results in terms of the regulation of gene activity by some kind of 'controlling elements' that had the ability to move from place to place on the chromosomes. As the elements move when the cells divide and proliferate, they mutate genes in only some of the cells — the changes in colour are due to their effects on pigmentation genes. McClintock's main conclusion was that these elements have a major influence on the development of organisms.

The discovery in the 1970s that bacteria, yeast

and fruitflies have DNA elements able to remove themselves and insert elsewhere in the genome prompted experiments that definitively identified transposable elements as major constituents of the genome. Somatic mutations (those that are not inherited) are typical of the action of TEs, but these elements also mutate genes in the germ line, leading to heritable genetic changes.

Interest in TEs increased further as drafts of the human genome became available from 2000, revealing that around 45% of the DNA consists of TEs. But even then, their role in gene regulation was not fully recognized. Today, however, TEs are acknowledged as a main component of most genomes, and McClintock's ideas that they can control gene activity are fully accepted. **C.B. & C.V.**

can affect gene regulation, these TE leftovers can contribute to genetic diversity even though the element is no longer present.

Not all TE excisions leave behind such an obvious footprint, however, so the influence of TEs on genomes has long been underestimated. But less noticeable TE remnants are being discovered, allowing the role of these elements in mutations and genome evolution to be elucidated. For instance, an ancient family of sequences that are derived from small interspersed nuclear elements (SINEs; Box 2) was discovered recently in the genomes of mammals⁴. Even though they do not encode proteins, these TE remnants seem to be under strong selective constraints, suggesting that they have some function that is being conserved during evolution.

The estimated rate of DNA mutation due to TE insertions differs between organisms. Among individual fruitflies, 50–80% of mutations are due to such insertions. This is a high value compared with that of 0.1–1% seen in the human genome, where the numerous copies of non-LTR retrotransposons are less active and mostly fixed in position. The reasons for this difference are not well understood, but global genomic characteristics, such as the rate at which genes are shuffled between generations (recombination), may be involved.

Expression

The expression of TEs — that is, the production of their encoded RNA — is tissue-specific; some elements are highly expressed during particular stages of the host organism's life, and some are even expressed differently in male and female reproductive cells (the germ lines). Such high levels and specific patterns of expression seem unexpected for apparently non-functional 'junk' DNA. This paradox had been explained by invoking either complex interactions between the TE regulatory sequences and the activity of numerous host developmental genes, or the influence of particularly highly expressed host genes on the TEs adjacent

to them (the 'readthrough phenomenon'). But both of these explanations implied a puzzling waste of energy for the cells and the organisms involved, so the idea that these RNAs are deliberately expressed to fulfil a particular cellular function gained ground.

This theory received a recent boost with the observation that some retrotransposons can influence the regulation of certain host genes and affect developmental processes in mouse oocytes (egg cells) and preimplantation embryos⁵. As well as validating McClintock's concept of TEs as controlling elements (Box 1), this finding reveals that TEs could have a role in the reorganization of genome structure and the gene silencing that occur during early embryonic development.

Transposable elements could have a marked impact on the relative contributions of the two sets of parental chromosomes in early development. In many animals, certain genes are differentially expressed according to their parental origin, and in mammals some TEs act rather like these 'imprinted' genes, in a manner dependent on their type. Thus, the DNA of SINEs is methylated — and thus inactivated — in the oocyte, but unmethylated in the male germ line. By contrast, intracisternal A particles (IAPs), retrotransposons and a long interspersed nuclear element (LINE) called LINE-1 display the opposite pattern⁶. This is consistent with the observation that maternally and paternally expressed imprinted genes in

the human genome tend to occupy distinct genomic regions that are activated differently in the male and female germ lines.

The silencing and activation of TEs in the male and female germ lines⁷ can affect genes that are involved in the RNA interference (RNAi) pathway. This is a process by which short, double-stranded RNAs induce the sequence-specific degradation of target messenger RNAs, halting production of the encoded protein. RNAi is thought to have evolved as a form of nucleic-acid-based immunity to inactivate viruses and TEs⁸.

In addition to their effects throughout embryonic development, TEs may act later in life. Indeed, LINE-1 retrotransposons seem to jump preferentially into the regulatory regions of certain neuronal genes in mice⁹. This alters the genes' expression, creating distinct populations of neuronal cells (somatic mosaicism). The impact of such TE activity on neuronal function remains to be fully elucidated, but if it occurs in humans, the activity of LINE-1 elements might be responsible for changes in the neuronal circuits in the brain, contributing to the intellectual diversity among people. This possibility opens an avenue of research that was previously unthinkable, even for the most imaginative believers in TEs as promoters of genetic and phenotypic diversity.

Environment

In many organisms, TEs are under 'epigenetic' control — where gene regulatory instructions are laid out in modifications of the DNA or its associated proteins (mainly histones) that do not alter the DNA sequence itself. These alterations include methylation of certain DNA nucleotides, and methylation or acetylation of the histone proteins around which the DNA is wound. Such epigenetic marks have a direct effect on the levels of expression of nearby genes (and TEs), and they are considered to be a second code in addition to the well-known code in the DNA sequence¹⁰.

Whereas some epigenetic instructions in the genome, such as gene imprinting, are usually long-lasting and heritable, the epigenetic code is more sensitive to environmental inputs than is the DNA sequence itself¹¹. Epigenetic modifications even diverge gradually over the lifetimes of monozygotic twins¹². Environmental stresses modify the genomic methylation state and thus change the structure of chromatin (the DNA and its associated proteins), inducing alterations in gene expression. Environmental stresses can thus markedly affect the way heritable changes in the DNA sequence (the genotype) are expressed as physical traits (phenotypes). Different stresses act differently on DNA methylation, histone methylation or acetylation, chromatin structure, and the production of small RNAs. Such processes interact strongly to affect genome function, and make up the core of epigenetic 'memory'¹³.

The influence of the environment on

Table 1 | Genome size and transposable elements

		Genome size (picograms)	% TEs
<i>Rana esculenta</i>	Frog	5.6–8.0	77
<i>Zea mays</i>	Maize	5.0	60
<i>Homo sapiens</i>	Human	3.5	45
<i>Mus musculus</i>	Mouse	3.4	40
<i>Drosophila melanogaster</i>	Fruitfly	0.18	15–22
<i>Caenorhabditis elegans</i>	Worm	0.1	12
<i>Saccharomyces cerevisiae</i>	Yeast	0.012	3–5
<i>Escherichia coli</i>	Bacteria	0.0046	0.3

gene expression can also be felt through effects involving TEs, because these elements can control genes epigenetically when inserted within or very close to them¹⁴. TEs can act either directly by inducing the methylation (and therefore silencing) of nearby DNA, or indirectly by disrupting the normal epigenetic state of a nearby gene. The control of coat colour by the *agouti* gene in mice is a classic example of such TE–gene interaction. Overexpression of the *agouti* gene leads to yellow fur, as well as to obesity, diabetes and an increased susceptibility to tumours. When an IAP retrotransposon inserts just in front of the *agouti* gene, the expression of the gene depends on the methylation state of the inserted IAP. Differences in the IAP methylation state in different cells leads to patchy coat colour. The methylation pattern of the IAP is completely erased in the male germ line but incompletely erased in the female germ line. So the variable expression of the *agouti* gene is passed on to subsequent generations as an epigenetic inheritance, mimicking a classical maternal environmental effect on offspring¹⁵.

The idea that environmental factors can affect how a phenotype is produced from a genotype

is reinforced by the observation that early nutrition can influence the expression of various genes, including TEs, at critical developmental stages¹⁶. The effect of nutrition on an organism's phenotype is often viewed as a manifestation of physiological misregulation, but it seems that changes in gene expression associated with TEs are involved as well. This should make us more conscious of the impact that our environment at one developmental stage can have on further stages, even in adults — especially if changes in TE expression under new environmental conditions do turn out to modify some neuronal functions, as proposed above.

Disease

The discovery that TEs can promote mutations and chromosome rearrangements, and can be activated by changes in epigenetic state, has led to the increasing realization that these elements may be responsible for various diseases — particularly cancers^{17,18}. Gene dysfunction or misregulation by TEs are considered to be responsible for around 0.5–1% of human illnesses¹⁹. Haemophilia, Duchenne muscular dystrophy, tumours of the oesophagus and reproductive organs and breast cancers may

result from the insertion of SINEs and LINEs near or within genes. The human genome has been colonized by thousands of copies of human endogenous retroviruses (HERV), which are related to retrotransposons, and these sequences are permanently integrated in the genome. Many HERV copies have been associated with teratocarcinoma and leukaemia, and others are involved in multiple sclerosis, schizophrenia and diabetes. But HERV insertions are not necessarily detrimental, and some copies have been co-opted by the genome to carry out cellular functions. For example, the syncytin gene corresponds to part of an endogenous retrovirus and is essential to the normal development of the placenta in humans and mice²⁰.

As the methylation states of TEs can be modified by environmental factors, it is possible that diet or chemicals that interfere with the DNA-methylating enzymes involved may have major effects on normal physiology and the manifestation of diseases such as cancers — effects that are as yet little understood. Moreover, caution should be exercised in the use of modified retrotransposons and retroviruses for delivering gene therapy — there has been

Box 2 | Structure and nomenclature of transposable elements

There is no officially agreed system for classifying transposable elements, but here is a simple system based on their evolution (phylogeny) and the genetic modules they contain (modified from ref. 30). The term transposon is often used as a generic term instead of transposable element, but it was originally coined to name the first characterized TE. This first TE moves about the genome through a DNA intermediate, using the transposase (Trp) enzyme to splice itself in and out of the DNA. The term DNA transposon (class II elements) is often used to distinguish this kind of TE from LTR (long terminal repeat) and non-LTR retrotransposons (class I elements) that move by means of an RNA intermediate and use the reverse transcriptase (RT) enzyme.

Inverted terminal repeats (ITRs) are needed for the movement of DNA transposons. The *gag* gene specifies the components of the molecular complex that is associated with the RNA transposition intermediate of retrotransposons. Retrotransposons also encode the RT enzyme, which synthesizes a complementary single-stranded RNA from the inserted DNA of the TE, and converts it to a double-stranded DNA that will be integrated in the genome elsewhere. They also encode the enzyme ribonuclease H (RH), which degrades the DNA–RNA hybrids obtained during transposition. Retrotransposons have genes encoding the integrase enzyme (INT), which splices the double-stranded DNA into a new spot in the host genome, and an enzyme (a protease; PR) that cuts up precursor proteins and is involved in particle assembly. Some retrotransposons have gained the envelope gene (*env*), which encodes surface proteins that

interact with the host cell membrane, conferring infectious characteristics on the elements.

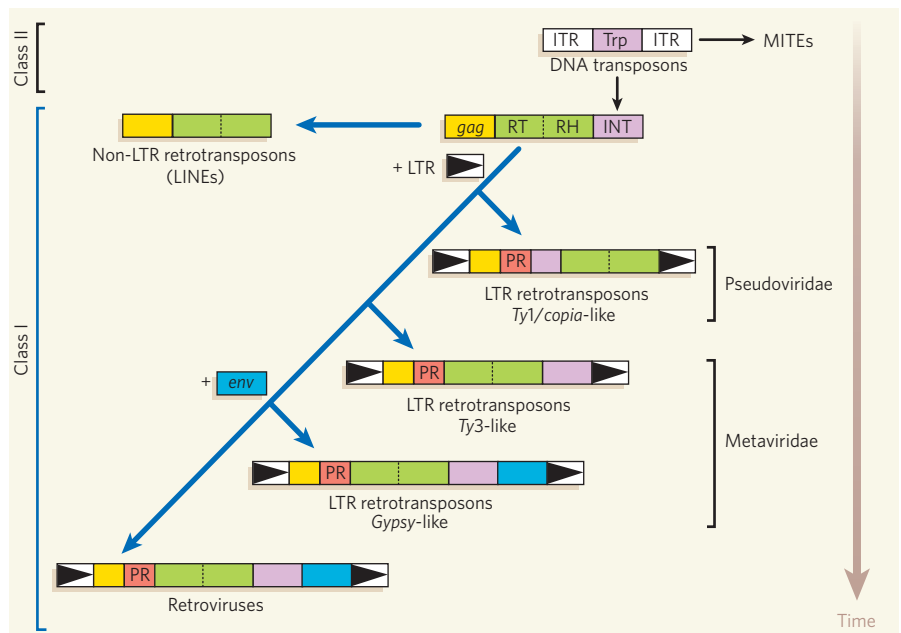
The human genome has about half a million copies of long interspersed nuclear elements (LINEs), of which 50–100 are still active. And there are more than a million copies of a short interspersed nuclear element (SINE) called *Alu* in the human genome. *Alu* elements are also a major source of genomic diversity in dogs. These SINEs have a particular structure and depend on LINEs for their transposition.

Miniature inverted-repeat transposable elements (MITEs) are an ancient TE family that

is characterized by short-sequence, terminal or subterminal inverted repeats flanked by short direct repeats, and they have no coding potential. They are distributed ubiquitously and seem to originate from DNA transposons.

Other classifications have been proposed on the basis of the transposition mechanism involved, for example, and a system similar to that used by virologists for viruses exists for LTR retrotransposons, which are then classified as Metaviridae and Pseudoviridae³⁰. But these classifications are not generally accepted.

C.B. & C.V.



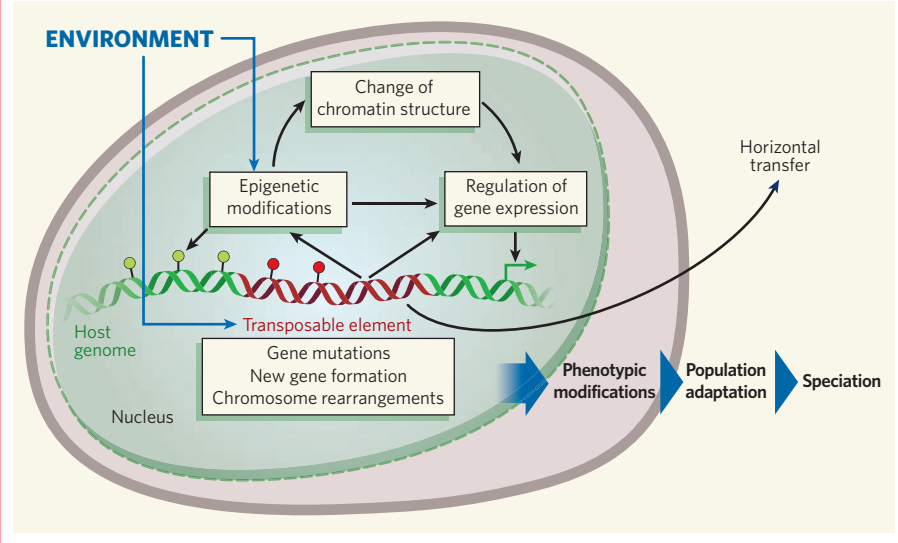
Box 3 | Transposable elements and genome–environment interactions

Transposable elements are increasingly seen as major originators of genetic change, allowing populations to adapt to change and species to evolve, as shown in the figure. They can also move between genomes of different species. Such horizontal transfer allows these elements to escape the various regulatory mechanisms imposed on them by their host genome, and to invade new genomes where they increase their copy number until new mechanisms evolve there to limit their spread.

Although many TEs are regulated in copy number and expression by various molecular mechanisms, some limiting forces are also at work at the population level. These forces

suggest that there is selection against the direct deleterious effects of insertions, even if these effects are small, and against the chromosomal rearrangements that frequently occur when TEs of the same family are present. The impact of such forces depends on the TE involved, the structure of the population and its reproductive system. It results in a slight tendency towards either the elimination of TE copies or their accumulation in genomes, making time an important factor in shaping the aspects of the genomes affected by TEs. As a result of these controlling forces, genomes contain a mixture of TEs, some of which are still active, whereas others are ancient relics that have degenerated.

C.B. & C.V.



an instance of retroviral activation of a cancer-promoting gene in gene-therapy trials²¹.

Populations

Because TEs can transpose at high frequency, with a rate ranging from 10^{-3} to 10^{-5} per element per generation, depending on the element, they are more powerful producers of the raw material of evolution than is the classical nucleotide-base substitution rate, which is around 10^{-8} – 10^{-9} per nucleotide per generation. So, waves of mobilization or loss through evolution may have had a major effect on the formation of new species, as has been suggested occurred in rodents²².

Species and populations differ in the structure and copy numbers of their TE sequences²³ (Table 1). Indeed, although the number of genes in a genome increases from bacteria to higher organisms, it is the proportion of TEs and other kinds of repeated sequence that accounts for the main differences in genome size among species. Genome size is thus not correlated with the complexity of the organism. From a population-genetics point of view, genome size may increase in a passive way as a consequence of small population size²⁴. This is because, in small populations, the efficiency of selection against the detrimental effects

of TE insertions is reduced and the effect of random processes is increased, leading to the accumulation of TEs. But it is an open question whether the variation in genome size is indirectly associated with host population size, or whether it is directly promoted by environmental stress or by the novel environmental conditions that populations encounter when they invade a new habitat. The answer will bear on our understanding of, for example, how ancestral humans adapted after they migrated out of Africa.

Although analysis of the genomes of various species is necessary to understand their variation in size, composition and structure, such an approach will give little information on the historical mechanisms that changed them. This knowledge will come from comparison of the genomes of many populations from a given species, together with their environmental conditions. For example, in the fruitfly *Drosophila melanogaster*, such work has shown that TE copy number varies greatly among populations, which have sometimes been invaded by novel TEs from other species (horizontal transfers)²⁵. Furthermore, waves of invasion or loss of TEs during evolution have been suggested for various elements, including the numerous LINES in the human genome, with some TEs

being in the process of invading while others are in the process of being eliminated²⁶. Population surveys should be able to test this theory, and to furnish material to analyse the forces that control TE and gene expression in nature and that shape our genome (Box 3).

Finding the links between changes in epigenetic gene regulation during the early and late stages of development on the one hand, and changes in the environment, genome size, and population size and structure on the other, is the next task for this field. Possible systems in which to study these effects have been uncovered recently. For example, variations in gene expression have been observed between genetically identical mice²⁷, together with variation in tissue-specific gene expression among natural populations of the teleost fish *Fundulus heteroclitus*²⁸. Also, strain-specific epigenetic variation that is associated with RNA abundance of a non-LTR retrotransposon has been found in the flowering plant *Arabidopsis*²⁹. Research into understanding these mechanisms and how they might apply to disease will then be crucial. Such research promises to shed light on environment–genotype interactions and on the link between genotype and phenotype. What was once dismissed as junk DNA must now be regarded as a major player in many of the processes that shape the genome and control the activity of its genes.

Christian Biémont and Cristina Vieira are in the Laboratoire de Biométrie et Biologie Evolutive, UMR 5558, CNRS, Université Claude Bernard Lyon 1, 69622 Villeurbanne Cedex, France. e-mail: biemont@biomserv.univ-lyon1.fr

1. Brandt, J. et al. *Gene* **345**, 101–111 (2005).
2. Medstrand, P. et al. *Cytogenet. Genome Res.* **110**, 342–352 (2005).
3. Messing, J. & Dooner, H. K. *Curr. Opin. Plant Biol.* **9**, 157–163 (2006).
4. Nishihara, H. et al. *Genome Res.* **16**, 864–874 (2006).
5. Peaston, A. et al. *Dev. Cell* **7**, 597–606 (2004).
6. Walter, J., Hutter, B., Khare, T. & Paulsen, M. *Cytogenet. Genome Res.* **113**, 109–115 (2006).
7. Kalmykova, A. I., Klenov, M. S. & Gvozdev, V. A. *Nucleic Acids Res.* **33**, 2052–2059 (2005).
8. Buchon, N. & Vauray, C. *Heredity* **96**, 195–202 (2006).
9. Muotri, A. R. et al. *Nature* **435**, 903–910 (2005).
10. Martienssen, R. A., Doerge, R. W. & Colot, V. *Chromosome Res.* **13**, 299–308 (2005).
11. McDonald, J. F., Matzke, M. A. & Matzke, A. J. *Cytogenet. Genome Res.* **110**, 242–249 (2005).
12. Fraga, M. F. et al. *Proc. Natl Acad. Sci. USA* **102**, 10604–10609 (2005).
13. Kim, S. Y. et al. *Plant Cell* **17**, 3301–3310 (2005).
14. Lippman, Z. et al. *Nature* **430**, 471–476 (2004).
15. Morgan, H. D. et al. *Nature Genet.* **23**, 314–318 (1999).
16. Waterland, R. A. & Jirtle, R. L. *Nutrition* **20**, 63–68 (2004).
17. Hughes, J. F. & Coffin, J. M. *Genetics* **171**, 1183–1194 (2005).
18. Feinberg, A. P. et al. *Nature Rev. Genet.* **7**, 21–33 (2006).
19. Kazanian, H. H. *Curr. Opin. Genet. Dev.* **8**, 343–350 (1998).
20. Frendo, J. L. et al. *Mol. Cell Biol.* **23**, 3566–3574 (2003).
21. Kaiser, J. *Science* **310**, 1894–1896 (2005).
22. Grahn, R. A. et al. *Cytogenet. Genome Res.* **110**, 407–415 (2005).
23. Biémont, C. & Vieira, C. *Cytogenet. Genome Res.* **110**, 25–34 (2005).
24. Lynch, M. & Conery, J. S. *Science* **302**, 1401–1404 (2003).
25. Kidwell, M. G. & Lisch, D. R. *Evolution* **55**, 1–24 (2001).
26. Fablet, M. et al. *Gene* **375**, 54–62 (2006).
27. Pritchard, C. et al. *Genome Biol.* **7**, R6 (2006).
28. Whitehead, A. & Crawford, D. L. *Genome Biol.* **6**, R13 (2005).
29. Rangwala, S. H. et al. *PLoS Genet.* **2**, 271–281 (2006).
30. Cappy, P. *Cytogenet. Genome Res.* **110**, 457–461 (2005).

BRIEF COMMUNICATIONS

Robotic whiskers used to sense features

Whiskers mimicking those of seals or rats might be useful for underwater tracking or tactile exploration.

Several species of terrestrial and marine mammals with whiskers (vibrissae) use them to sense and navigate in their environment — for example, rats use their whiskers to discern the features of objects¹, and seals rely on theirs to track the hydrodynamic trails of their prey². Here we show that the bending moment — sometimes referred to as torque — at the whisker base can be used to generate three-dimensional spatial representations of the environment, and we use this principle to construct robotic whisker arrays that extract precise information about object shape and fluid flow. Our results will contribute to the development of versatile tactile-sensing systems for robotic applications, and demonstrate the value of hardware models in understanding how sensing mechanisms and movement control strategies are interlocked.

Rats actively whisk (rotate) their whiskers against objects during exploration¹, whereas seals keep their whiskers relatively fixed in order to track fluid wakes². Using classical elasticity theory^{3,4}, we modelled the rat whisker as a conical beam bending against an object, and derived a monotonic relationship between the radial contact distance, d , and rate of change of moment, M , at the whisker base (see supplementary information). This theoretical solution was experimentally verified by using a strain gauge to measure moment at the base of a rat's plucked whisker as it was smoothly rotated against a slender peg placed at various radial distances (Fig. 1a). Agreement was excellent between the actual radial distance and that obtained from the experiment (for details, see supplementary information).

To investigate the shape-extraction capability of this encoding mechanism, we constructed a 4×1 array of robotic whiskers. Each spring-steel wire whisker was fitted with four strain gauges at its base to measure the two orthogonal components of moment. We tested the array by whisking it across a small, sculpted head, an object selected specifically for its intricate concavities and convexities (for details and video, see supplementary information). Whisking was performed with the array at evenly spaced heights and angles. Analysis of moments on each whisk provided four radial contact points in three-dimensional space (one for each whisker). Fitting a smooth surface to the contact points then resulted in a faithful extraction of the original shape of the sculpture (Fig. 1b).

To test the ability of whiskers to extract the

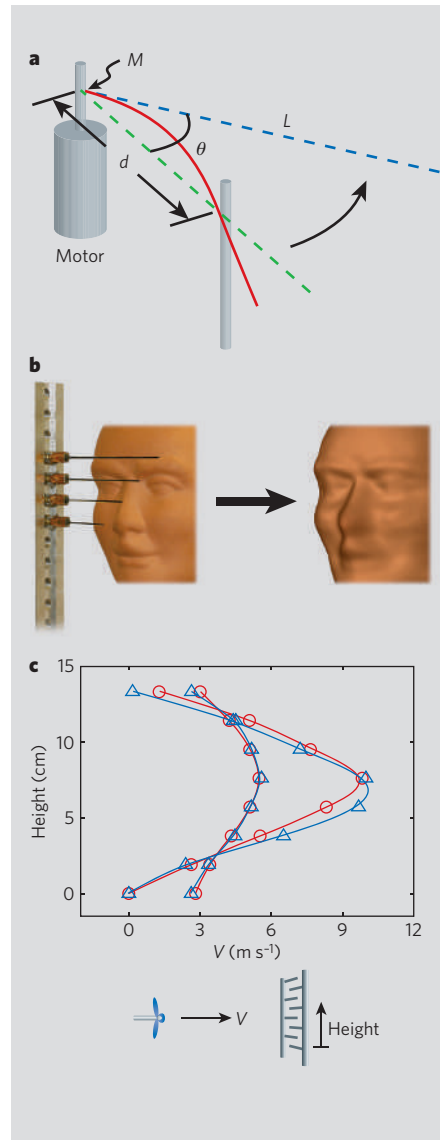


Figure 1 | Whisker arrays extract information about complex environmental features. **a**, A whisker (red) of length L rotates through an angle θ against an object at distance d , which elicits a moment M at the base. **b**, Extraction of information about an object shape using an artificial whisker array (whisker lengths are 2, 3, 4 and 5 cm). **c**, Fluid-flow profiles, determined using two opposing whisker arrays, arranged as shown in the inset. V is the fluid velocity and each whisker is $0.5 \text{ cm} \times 11 \text{ cm}$. Blue curve: data from whisker array; red: data from a Pitot tube placed at the height of each whisker, which gives the true air speed. Standard deviations (10 trials) are within the size of the symbols.

features of fluid flows, we used two opposing 4×1 whisker arrays (Fig. 1c, inset). Each stainless steel wire was replaced with a thin, flexible, plastic strip to ensure that there was a large surface area normal to the flow and to maximize bending. When a stream of air moving at velocity V was directed towards the centre of the array, each whisker was deflected by a magnitude that was dependent on its distributed load, permitting an accurate characterization of the shape of the fluid stream (Fig. 1c). More complex whisker configurations would allow flow patterns to be described in three dimensions, which might enable moving underwater objects to be tracked by wake.

Until now, neural responses of the whisker system have been described mainly in terms of kinematic variables, that is, position and its variation with time. Our findings indicate that the primary sensory neurons of animals that use their whiskers for exploration may also encode information about moment. This idea is consistent with physiological recordings from the rat trigeminal ganglion⁵. Sensing moment at the whisker base during contact with an object could be a low-frequency counterpart to the whisker vibrations that are thought to be used to discriminate textures^{6,7}.

Our results on biomimetically engineered whiskers may find application in land-based robots and autonomous underwater vehicles, in which a capability for tactile perception could broaden and enhance performance.

Joseph H. Solomon*, **Mitra J. Hartmann*†**

Departments of *Mechanical Engineering and †Biomedical Engineering, Northwestern University, Evanston, Illinois 60208, USA
e-mail: m-hartmann@northwestern.edu

1. Krupa, D. J., Matell, M. S., Brisben, A. J., Oliveira, L. M. & Nicolelis, M. A. L. *J. Neurosci.* **21**, 5752–5763 (2001).
2. Dehnhardt, G., Mauck, B. & Bleckmann, H. *Nature* **394**, 235–236 (1998).
3. Kaneko, M., Kanayama, N. & Tsuji, T. *IEEE Trans. Robot. Autom.* **14**, 278–291 (1998).
4. Clements, T. N. & Rahn, C. D. *IEEE Trans. Robot.* **22**, 844–848 (2006).
5. Szwed, M. et al. *J. Neurophysiol.* **95**, 791–802 (2006).
6. Neimark, M. A., Andermann, M. L., Hopfield, J. J. & Moore, C. I. *J. Neurosci.* **23**, 6499–6509 (2003).
7. Hartmann, M. J., Johnson, N. J., Towal, R. B. & Assad, C. *J. Neurosci.* **23**, 6510–6519 (2003).

Supplementary information accompanies this communication on Nature's website.
Received 30 June; accepted 8 September 2006.
Competing financial interests: declared none.
doi:10.1038/443525a

BRIEF COMMUNICATIONS ARISING online
www.nature.com/bca see Nature contents.

GLUCONEOGENESIS

Re-evaluating the FOXO1-PGC-1 α connectionArising from: P. Puigserver *et al.* *Nature* 423, 550–555 (2003).

Increased expression of the gene encoding the enzyme glucose-6-phosphatase (G6Pase) contributes to the increased production of glucose by the liver that occurs in individuals with diabetes. Puigserver *et al.*¹ show that the transcription factor FOXO1 and the transcriptional co-activator PGC-1 α act synergistically to stimulate the expression of genes in the gluconeogenesis pathway and propose that PGC-1 α acts, in part, directly through FOXO1. Here we show that FOXO1 is neither required nor sufficient for the stimulation of G6Pase–luciferase fusion gene expression by PGC-1 α . Our results indicate that the transcriptional interaction between FOXO1 and PGC-1 α is indirect.

PGC-1 α (for peroxisome proliferative activated receptor- γ co-activator-1 α) interacts with several transcription factors and is an important regulator of mitochondrial biogenesis, respiration, thermogenesis and hepatic gluconeogenesis^{2,3}. Puigserver *et al.*¹ show that FOXO1 and PGC-1 α can physically interact with one another and that the combined action of PGC-1 α and FOXO1 in various liver-cell types results in a synergistic induction of endogenous G6Pase gene expression. They also show that, in immortalized hepatocytes, mutation of FOXO1 binding sites in the G6Pase promoter abolishes the induction of fusion-gene expression by FOXO1 alone, as well as the synergistic induction achieved in combination with PGC-1 α (ref. 1).

On the basis of this result, the authors propose a model in which PGC-1 α stimulates G6Pase gene expression, in part, through direct interaction with FOXO1 bound to the G6Pase promoter. However, an alternative explanation for their observations is that PGC-1 α stimulates G6Pase gene expression through interaction with another promoter-bound transcription factor, distinct from FOXO1, and that the synergy between PGC-1 α and FOXO1 occurs through interaction with the pre-initiation transcription complex. It was not easy for Puigserver *et al.*¹ to distinguish between these possibilities in immortalized hepatocytes because there was little effect of PGC-1 α in the absence of FOXO1.

This is not the case in H4IIE hepatoma cells, however, in which both PGC-1 α (ref. 4) and FOXO1 (ref. 5) independently induce G6Pase–luciferase fusion gene expression, although, as in immortalized hepatocytes, co-expression of both PGC-1 α and FOXO1 results in a synergistic induction of fusion-gene expression (results not shown). We therefore generated plasmids in which two previously identified FOXO1 binding sites (designated IRS 1 and IRS 2; ref. 6) and two

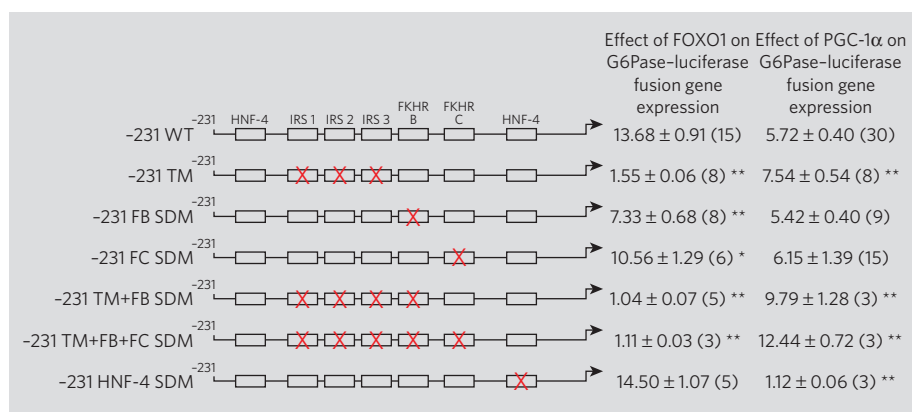


Figure 1 | Mutation of FOXO1 binding sites (IRSs 1, 2, 3 and FKHR B and C motifs) in the promoter of the gene encoding glucose-6-phosphatase (G6Pase) abolishes the induction of G6Pase fusion-gene expression by FOXO1 but not by PGC-1 α . H4IIE cells were transiently co-transfected with various G6Pase–luciferase fusion genes (12 μ g) and expression vectors encoding *Renilla* luciferase (0.15 μ g) and one of pBJ5 (3 μ g), pBJ5-PGC-1 α (3 μ g), pcDNA3 (3 μ g) or pcDNA3-FOXO1 (3 μ g). The G6Pase–luciferase fusion genes incorporated either the wild-type (WT) promoter sequence, located between –231 and +66, or contained the same promoter fragment with site-directed mutations (SDMs) in the three IRS motifs (TM), the FKHR B (FB) and C (FC) motifs, and the HNF-4 binding site in various combinations, as indicated. The IRS 3 motif does not bind FOXO1 but its sequence is similar to that of IRS 1 and IRS 2 (ref. 6). After transfection, cells were incubated for 18–20 hours in serum-free medium, then collected and assayed for luciferase expression. Results are shown as the ratio of firefly luciferase activity, corrected for *Renilla* luciferase activity in the cell lysate, in FOXO1- and/or PGC-1 α -stimulated versus control cells, expressed as fold induction; values are shown as mean $\pm$ s.e.m. for the indicated number of experiments, using at least three independent preparations of each luciferase fusion-gene plasmid and duplicate assays for each sample. Double asterisks, $P < 0.05$ versus WT; single asterisk, $P = 0.05$ versus WT for FC SDM.

recently identified FOXO1 binding sites (designated FKHR B and C; ref. 5) in the G6Pase promoter were mutated. We then investigated the effect of these mutations on the ability of PGC-1 α to stimulate G6Pase fusion gene expression.

Figure 1 shows that mutation of these FOXO1 binding sites abolishes the induction of fusion-gene expression that is seen with FOXO1 alone. In contrast, the mutations do not reduce the induction of fusion-gene expression by PGC-1 α , but actually increase the effect (Fig. 1). These results indicate that the effect of PGC-1 α does not require interaction with FOXO1 bound to the G6Pase promoter, although PGC-1 α can synergize with FOXO1 to stimulate G6Pase fusion gene expression (ref. 1, and data not shown).

On the basis of a 5'-deletion analysis, we previously identified a hepatocyte nuclear factor-4 α (HNF-4 α) binding site in the mouse G6Pase promoter, located between nucleotides –76 and –64, that seemed to mediate PGC-1 α -stimulated G6Pase fusion gene expression⁴. A caveat for this approach is that the role for this HNF-4 α binding site was only demonstrated in the context of a promoter fragment in which all four FOXO1 binding sites had been deleted. We therefore considered the possibility of

redundancy, whereby PGC-1 α -stimulated G6Pase fusion gene expression could be mediated either through the HNF-4 α binding site or through the FOXO1 binding sites.

To investigate this possibility, a mutated G6Pase promoter was generated that contained a site-directed mutation in the HNF-4 α binding site but that had all four FOXO1 binding sites left intact. This mutation had no effect on FOXO1-stimulated fusion-gene expression (Fig. 1), whereas the induction of fusion-gene expression by PGC-1 α was abolished, together with the synergistic effect of PGC-1 α and FOXO1 (Fig. 1, and results not shown).

These results show that the HNF-4 α binding site is required for the PGC-1 α response and that the binding of FOXO1 to the G6Pase promoter is neither required nor sufficient for the stimulation of G6Pase fusion gene expression by PGC-1 α . This conclusion is consistent with the observation that the ability of PGC-1 α to stimulate G6Pase expression is completely lost in hepatocytes derived from mice lacking HNF-4 α (ref. 7).

Marcia M. Schilling, James K. Oeser, Jared N. Boustead, Brian P. Flemming, Richard M. O'Brien

Department of Molecular Physiology and

Biophysics, Vanderbilt University Medical School,
Nashville, Tennessee 37232, USA
e-mail: richard.obrien@vanderbilt.edu

1. Puigserver, P. *et al.* *Nature* **423**, 550–555 (2003).
2. Finck, B. N. & Kelly, D. P. *J. Clin. Invest.* **116**, 615–622 (2006).
3. Lin, J., Handschin, C. & Spiegelman, B. M. *Cell Metab.* **1**, 361–370 (2005).

4. Boustead, J. N. *et al.* *Biochem. J.* **369**, 17–22 (2003).
5. Vander Kooi, B. T. *et al.* *Mol. Endocrinol.* **19**, 3001–3022 (2005).
6. Vander Kooi, B. T. *et al.* *J. Biol. Chem.* **278**, 11782–11793 (2003).
7. Rhee, J. *et al.* *Proc. Natl Acad. Sci. USA* **100**, 4012–4017 (2003).

doi:10.1038/nature05288

REVIEWS

From *in vivo* to *in silico* biology and back

Barbara Di Ventura^{1*}, Caroline Lemerle^{1*}, Konstantinos Michalodimitrakis¹ & Luis Serrano¹

The massive acquisition of data in molecular and cellular biology has led to the renaissance of an old topic: simulations of biological systems. Simulations, increasingly paired with experiments, are being successfully and routinely used by computational biologists to understand and predict the quantitative behaviour of complex systems, and to drive new experiments. Nevertheless, many experimentalists still consider simulations an esoteric discipline only for initiates. Suspicion towards simulations should dissipate as the limitations and advantages of their application are better appreciated, opening the door to their permanent adoption in everyday research.

Intuition and concepts constitute, therefore, the elements of all our knowledge, so that neither concepts without an intuition in some way corresponding to them, nor intuition without concepts, can yield knowledge... Thoughts without content are empty, intuitions without concepts are blind... Only through their union can knowledge arise.¹

In the past few years, the biological community has been exposed to a new buzzword: 'systems biology'. Irrespective of its exact definition, there are two attributes of systems biology on which most agree: an '-omics' aspect (involving the comprehensive collection of experimental data concerning a system), and the use of mathematical modelling to make testable predictions and gain insight about a biological system's behaviour. The intrusion of computational biology into 'wet' laboratories is producing a quiet revolution wherein simulation tools are used to complement experiments and accelerate the hypothesis generation and validation cycle of research^{2–4}. Modelling a cellular process can highlight which experiments are likely to be the most informative in testing model hypotheses, and allow testing for the effect of drugs⁵ or mutant phenotypes⁶ on cellular processes—thus paving the way for individualized medicine⁷.

Simulation of biological systems is a broad field, and the focus of this Review is mainly on networks operating at the cellular level^{8,9}. Here, we expose the problems and limitations of models and simulations, and suggest when and how to use them. Our aim is to familiarize experimentalists with this exciting multidisciplinary endeavour so they might consider complementing their current tools with computational ones. As recent progress in the field has shown, computational biology can successfully assist experimentalists in unravelling the principles and operation of complex systems.

Modelling and simulation of biological systems

Biologists commonly use the term 'model' for verbal or graphical descriptions of a mechanism underlying a cellular process. Less often do they use it to refer to a set of equations expressing in a formal and exact manner the relations among variables that characterize the state of a biological system. The approach of biologists towards knowledge building has been mostly empirical (following 'intuition'), but experimental facts remain 'blind' without laws or principles derived from them. Conversely, theoretical approaches used by modellers have often failed to relate to real systems, such that theoretical concepts encapsulated in these studies are equally 'empty'. Instead, theory and experiments need to be viewed in close interplay.

Mathematical models are more rigorous and powerful than descriptive ones. In some cases, concepts derived from engineering—

like the dampening effect of negative feedback on noise¹⁰—apply directly even to graphical models, but most often they provide nothing more than a mere indication of how a system might behave¹¹.

Bridging the gap between a large body of experimental data and a potentially useful mathematical model is not trivial. Knowledge about the system is essential and needs to be formalized for the chosen framework: for instance, by breaking reactions into mono- and bi-molecular ones if only elementary reactions are allowed. Ideally, all information relevant to a system (not only concentrations and rates of events, but also spatial distribution, diffusion parameters, excluded volumes, and so on) would be known to make a maximally accurate *in silico* replica of the system. Unfortunately, even for the best-studied systems, the mass of accumulated data still falls short of describing, even qualitatively, the variety of elementary processes that each molecular species engages in (post-translational modifications, degradation, complex formation, and so on); even less known are details of spatial information and the timing of events. Consequently, assumptions are necessary (for example, that all gene copies of a multi-copy plasmid are transcriptionally active, or that a certain molecule freely diffuses inside a cell or is always monomeric). On the other hand, it can be beneficial to exclude some known data to accommodate available computational power and to facilitate the analysis (even at the expense of accuracy). For example, irrelevant interactions of highly connected proteins could be omitted; details such as the cell-cycle regulation of a certain protein could be temporarily set aside; abundant species such as ATP or ribosomes might be represented as constant pools; or transcription and translation events might be lumped together.

With quantitative information often being both incomplete and of non-uniform quality, many modelling tools that are tailor-made to qualitative data are available (in addition to sometimes better known quantitative modelling approaches), some of which also give rise to quantitative predictions (for example, constraint-based modelling approaches). Given the limitations of any single modelling approach (mathematical tools applicable, data types allowed, and so on), it is often worthwhile trying a combination of them¹², and there are indeed a variety of diverse formalisms to choose from including some increasingly promising ones, such as Petri nets¹³ and concurrency-theory-derived methods¹⁴, inherited from computer science.

Qualitative modelling

The majority of current experimental techniques, including high-throughput ones, yield only qualitative or semiquantitative data. For

¹European Molecular Biology Laboratory, Meyerhofstrasse 1, 69117 Heidelberg, Germany.

*These authors contributed equally to this work.

simulations tools to be applied and useful in drawing non-obvious conclusions, the analysis of such data needs to allow, as a minimum, the formulation of logical statements describing, for instance, causal relationships between events involving model components. As an example, computer science algorithms used to perform code checks can assess the logical consistency of a set of statements: that is, check that no subset of statements is in contradiction with any other¹⁵. Automated tools such as these and others used in qualitative reasoning approaches become indispensable if logical inferences are to be made on very large sets of experimental observations. In qualitative modelling, kinetic processes¹⁶ are simulated by tracking over discrete time the state of the system, defined in terms of a coarse range for each variable. The weak specification of such models conserves computer resources needed to explore the space of possible behaviours; moreover, it provides high-level predictions applying to a whole family of systems—for instance, the number of feedback loops or the ranges of variables supporting oscillations or switches. Although simulation of qualitative models can be fast, even a rough exploration of parameter space can become intractable as the size of the system increases, highlighting the need for increasing computer resources and methods to accelerate the parameters search.

In so far as biological systems have evolved tolerance to random fluctuations and perturbations, coarse ranges may suffice to predict correctly a system's behaviour¹⁷. For genes that are naturally found in only two states¹⁸, the trade-off in accuracy may not even be high. On the other hand, simple models can, in some cases, predict behaviours that are far away from reality (Fig. 1 and below).

Constraint-based modelling

Many biological systems assume dynamic steady states, with the environment acting as source and sink for some molecules, whereas the concentration of the other molecules is balanced by the activities of the reactions (reaction fluxes) in the system. These states depend as much on specific properties of the components (for example, stoichiometries and reversibility of the reactions, enzymatic capacities) as they do on general constraints (for example, the law

of mass action and laws of thermodynamics). Constraint-based modelling (Supplementary Fig. 1) incorporates this information in a framework that allows one to identify all sets of reaction fluxes that achieve steady state for a reaction network of interest. Built on precursor fields such as pathway analysis¹⁹, this approach uses a rich variety of rigorous mathematical and computing tools (linear algebra, convex analysis and linear programming, to name but a few) to achieve results of high predictive value²⁰ along two lines of analysis: network-based pathway analysis (to extract systemic properties of the network such as pathway redundancy²¹, missing reactions²², and so on) and flux balance analysis²³ (to identify those system states in which so-called objective functions—for example, growth rate and ATP consumption—reach an optimal value). Adding constraints increases the accuracy of predictions by reducing the number of allowable states—for instance, by specifying conditions when reaction fluxes are zero because the corresponding enzymes are not expressed²⁴. When a network's response to an environmental change is very fast compared with that change (that is, when the quasi-steady-state assumption holds) its time-evolution can be calculated recursively as a series of steady states (dynamic flux balance analysis; see Supplementary Fig. 1d). Having to consign reactions in a stoichiometry matrix, imposing the separation of substrates and products (for example, metabolites) from reaction effectors (for example, enzymes), means constraint-based modelling is rather inflexible; in particular, transcription regulatory gene networks and signal transduction cascades do not fit the formalism well. In fact, for them—as the quasi-steady state assumption is generally not valid and transient responses, inaccessible using this approach, are often the point of interest—other formalisms are often preferred.

Beyond the limited scope of application and the strong focus on steady states, a drawback of this method is the heavy computational load required to calculate optimal points, and its Achilles' heel is the choice of objective function²⁵. Nevertheless, this technique has proved to be well-suited to analyse complex steady states and system responses to perturbations, whether genetic or environmental.

Quantitative modelling

Compared with qualitative models, quantitative ones have a natural appeal in that they offer greater detail in mimicking reality. Moreover, rich qualitative insights on the system are possible using theoretical tools such as bifurcation and stability analysis²⁶, which, for example, indicate the precise boundaries of parameter ranges to which steady states or sustained oscillations correspond, or reveal the stability of the solutions before actually solving the dynamical equations representing the system.

Quantitative models can be either deterministic or stochastic. The most popular formalism is the deterministic ordinary differential equations (ODEs) one which, when extended to model space, is referred to as partial differential equations (PDEs). Each equation in a set typically represents the rate of change of a species' continuous concentration as a sum or product of, more or less, empirical terms (typically law of mass action terms, Michaelis–Menten functions, power laws²⁷, and so on), accounting for the effect of biological events on such concentrations. By definition, the initial state of the system in a deterministic model uniquely sets all future states. As analytical solutions seldom exist, numerical solutions need then to be computed (once for each set of parameter values and initial conditions explored). A word of caution: although this step is simple in principle, wrong solutions can arise. For instance, the chosen step-length for the integration of the ODEs can be sufficiently large to cause divergence of the numerical solution from the correct one (numerical instability), making a minimum of experience with related issues a strong asset for the user. In general, ODEs are best suited to capturing the behaviour of systems where species are abundant and reaction events frequent (as is often the case for metabolic pathways, for example), because species concentrations are then acceptably approximated as varying continuously and predictably.

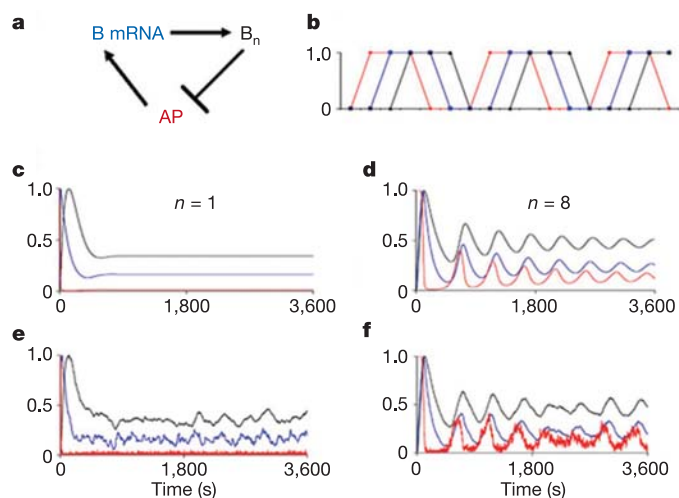


Figure 1 | Simulation of a simple network using different mathematical formalisms. **a**, Diagram of the negative feedback network used in the simulations. n , the number of B molecules in the active complex. **b–f**, Time courses of activator protein AP (red), B mRNA (blue) and B protein (black). The y axis represents the number of molecules, normalized for each species by the maximum value reached, except in **b**, in which it represents presence or absence of the molecules. Simulation of discrete time boolean model (**b**) with synchronous update. Deterministic (**c**, **d**) and stochastic (**e**, **f**) simulations using specified parameters (see Supplementary Fig. 2), with B monomer (**c**, **e**) or octamer (**d**, **f**). Oscillations predicted by the boolean model are obtained in the deterministic/stochastic model only when B oligomerization is included.

Molecular interactions are intrinsically random and cellular behaviour itself sometimes seems to reflect this randomness²⁸. Indeed, occurrences of noise have been found to be exploited by cells—for instance, to survive a variety of environmental changes²⁹ or to increase sensitivity in signal transduction processes³⁰. To model such stochastic systems, two main methods are used. The first comprises using stochastic differential equations (SDEs; derived from ODEs by adding noise terms to the equations), the solutions for which can be numerically obtained either by computing many trajectories (Monte Carlo methods) or approximating their probability distribution and then calculating statistical measures (such as mean and variance). Notably, with this method noise is imposed on the system and represented by mathematical terms chosen *a priori*, instead of arising from the underlying physical interactions. The second is a very successful and exact method introduced nearly 30 years ago^{31,32}, and recently enhanced to cope with different reaction timescales^{33–35} or space^{36–38}. With this approach, molecules are modelled individually and reaction events are calculated by their probability. The price to pay for having a more physically realistic model is the considerable increase in computational time and the need for specialized algorithms^{36,37}.

At the interface between purely deterministic and stochastic modelling techniques, novel hybrid approaches have appeared that speed up simulations considerably by partitioning reactions into fast reactions, assigned to a continuous framework, and slow reactions, simulated with a discrete stochastic algorithm^{39,40}.

Space in modelling

Until recently, the majority of simulations ignored the fact that biological processes take place in heterogeneous and highly structured environments. Even prokaryotes are now known to possess a cytoskeleton and control the movement and location of molecules, so regulating cellular processes in both space and time⁴¹. Indeed, spatial segregation underlies many cellular strategies; reactions are prevented by physically separating molecules, and molecular gradients⁴² within or between cells are used in pattern formation.

Crucial as it may be for fundamental processes such as self-organization^{43,44}, morphogenesis⁴⁵, cell division⁹ or calcium waves⁴⁶, spatial information is still largely absent from interaction databases. Recent technological advances are addressing this dearth of spatial data, and theoretical advances are improving computational methods, making it now possible to simulate spatio-temporal models of biological processes in coarse-grained or realistic geometries⁴⁷.

When to use different models: problems and solutions

Modelling could seem simple to an outsider: define your system, choose a modelling approach on the basis of what you know and want to know, download a simulation tool, input some parameters, run the simulation, and collect the results. However, as in the earlier days of protein design, what looks nice on a screen does not necessarily carry any biological meaning. There are many conceptual pitfalls for the modeller, which result in unrealistic predictions.

Mathematical-formalism-independent errors. Independent of the mathematical formalism used, many obvious errors can be introduced in a model. One set of errors arises from using parameter values measured in the often dilute, homogeneous environment of a test tube to model the crowded, apparently messy environment of cells. For example, the apparent binding constant of two interacting domains *in vivo* may differ from that obtained *in vitro* owing to crowding effects and/or spatial constraints⁴⁸. Similarly, bulk estimates of the number of molecules per cell may hide local concentration variations that would affect reactions rates. Near the cell membrane, where the negative potential changes the electrostatic component, k_{on} values can be affected. Cell line, host and experimental conditions bring variability to data collected *in vivo* that needs to be considered.

Other subtle details can, if overlooked, lead to completely wrong

predictions. For instance, assuming that the degradation rate of a protein is the same whether alone or in a complex is often unjustified, because complex formation can readily affect protein stability. In general, the common practice of assigning ‘default’ values to generic reactions (degradation rates, transcription activity, and so on) is acceptable only as a first approach and begs further refinement. The precise order in which a series of reactions occurs can be consequential, yet is frequently disregarded. For example (Fig. 2), if the formation of active dimers of a certain species is known to trigger downstream events, and, as an additional assumption, if such a dimer is only active when both monomers are activated (for example, by phosphorylation), one could formulate a general model considering all possible dimerization and monomer activation/deactivation reactions derived from these hypotheses. Not knowing the order of events, three extreme possibilities can be formulated: two of total dependence (only activated monomers can dimerize or dimerization must occur before activation can) and one of total independence between activation and dimerization reactions, leading to different steady-state concentrations of active dimer. Finally, when running a simple simulation over many cellular generations, surprises might be avoided by remembering to consider the effect of cell division on molecular pools.

Mathematical-formalism-dependent errors. Problems can arise from the mathematical formalism used to simulate a system. To illustrate the impact of modelling choices, simulations of a simple gene network with negative feedback (protein B forms multimers and sequesters the activator protein AP responsible for its transcription) were run using three formalisms with different degrees of graininess (Fig. 1): a simple boolean model, a quantitative deterministic model with ODEs, and a quantitative stochastic model³⁷. With the discrete-time boolean model, the built-in delay produces oscillations (Fig. 1b).

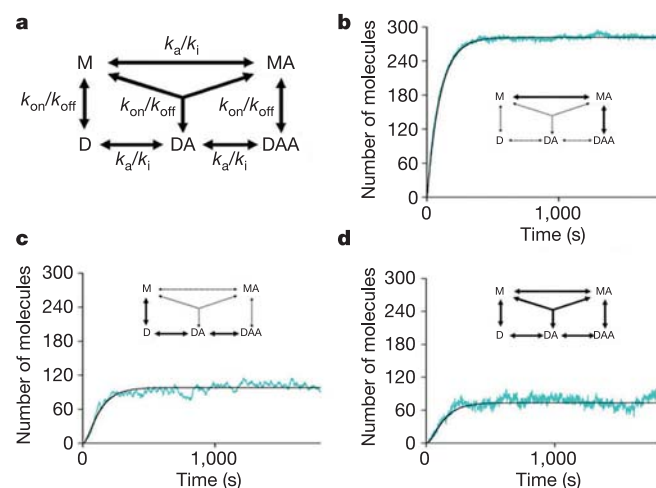


Figure 2 | Example of mathematical-formalism-independent pitfalls in modelling. If protein M is active as a post-translationally modified dimer DAA, the formation of DAA can be modelled differently, depending on assumptions made regarding the order of events and the nature of the active dimer. Here we assume that both monomers must be modified to form an active dimer. **a**, The scheme shows all reactions compatible with the experimental observations. M, monomer; MA, modified monomer; D, dimer; DA, dimer with one modified monomer; DAA, active dimer (both monomers modified). Each horizontal arrow corresponds to a reversible activation reaction and each vertical arrow corresponds to a reversible dimerization reaction. The two slanted arrows, together with the central vertical arrow, represent dimerization of a modified monomer with an unmodified one. **b–d**, Simulation runs for three different orders of events—respectively, ‘activate then dimerize’, ‘dimerize then activate’ and ‘dimerize and activate together independently’—showing the deterministic (black) and stochastic (cyan) temporal evolution of DAA molecules (see Supplementary Fig. 3 for parameters used in the simulation and a mathematical derivation of the steady-state solution).

The other two models require additional events to be modelled explicitly (for example, degradation to balance production), and in contrast to what is observed with the coarser boolean model, oscillations did not occur unless multimerization was allowed (Fig. 1c–f).

Accounting for constraints of cellular space can significantly impact predictions, making spatial models more powerful compared with non-spatial ones. In Fig. 3, two different behaviours of a simple network are modelled, in which a phosphorylated transcription factor triggers the production, in one pole of the cell, of protein A that is involved in a positive feedback loop (resulting from mutual repression of species A and B). Starting from an initial state corresponding to a high concentration of B and no A, when the kinase and the phosphatase freely diffuse in the cell, the positive feedback acts as a switch causing the disappearance of B and the accumulation of A (Fig. 3d). When the kinase and the phosphatase are localized to opposite poles of the cell (Fig. 3b), however, a gradient of the phosphorylated transcription factor is formed (Fig. 3c), and the amount of A produced is insufficient to trigger the switch (Fig. 3e).

Because molecules are discrete species, continuous representations of molecular abundance are another source of artefacts. For example, contrary to what happens with discrete-valued models, steady states take infinite time to be reached with continuous concentrations, a discrepancy that disappears by focusing instead on how fast the steady state is approached (that is, by introducing the concepts of half-life, rise-time and others; Fig. 4a). Similarly, a probabilistic key is needed to interpret the non-integral values generated by the continuous-value deterministic models, as for discrete stochastic models. For low numbers, however, this interpretation is not error proof. If we consider a single, strictly autocatalytic species, (Fig. 4b), an ODE model would conclude (as would a related SDE model with a noise term that was only multiplicative) that there are two steady states: an unstable lower one (with zero molecules per cell) and a stable upper one. Simulation runs based on such models show that any non-zero initial state evolves asymptotically towards the upper steady state, whereas zero states are absorbing (that is, that once reached, they cannot be left). If important downstream events, such as cell differentiation or apoptosis, are triggered only by the absence of B molecules, a continuous model would lead us to wrongly conclude that these events never occur, whereas the more physically realistic, discrete stochastic model would reveal that, for each cell, it is

only a matter of time before the triggering state is reached (Fig. 4b).

In summary, the modeller should never draw conclusions from simulation results without keeping in mind the limitations of a given approach to represent reality.

Consideration of the biological question. In the fortunate case in which both qualitative and quantitative information are available for our system and computational power is not limiting, the choice of formalism depends on which is best suited to capture the essential properties of the real system. Should time and variables in the model have continuous or discrete values? Does localization play an important part in any of the reactions? If so, which model geometry should be considered? Should fluctuations be allowed? Something that greatly helps in this decision process is having a clear biological question to answer and some ideas about the functional phenomenon being modelled. Thus, if the system includes gradients, the mathematical formalism used should handle space. If the system seems noisy (for example, not all cells respond in the same way to the same stimulus), then a stochastic approach might clarify this point. For metabolic pathways especially, constraint-based modelling provides a wealth of avenues to explore. Although these examples may induce wariness in the novice modeller when trying to decide on an approach, such considerations will become second nature as more practice in the 'systems biology' mentality is gained.

Successes in modelling

A model should allow the behaviour of a system to be predicted under various conditions, including some not yet experimentally tested and therefore not exploited in the model-building process. This most often requires an iterative process of model building and experimental investigation. Although most current models are at an early stage of this process, the number of cases where the symbiotic interaction of experiment and model has led to successful predictions is steadily growing. The first examples came in the field of metabolic engineering, where modelling led to the identification of bottlenecks in production⁴⁹. More recently, as constraint-based modelling approaches are applied to reconstructed networks of genome scale, significant progress is being made in understanding and predicting the phenotypic potential of microorganisms^{50,51}.

Models have been most successful for systems of simpler organisms, like *Escherichia coli* and *Saccharomyces cerevisiae*. Two interesting examples are those of sphingolipid metabolism⁵² and the osmotic shock response⁵³ in *S. cerevisiae*, with the latter being a good example of

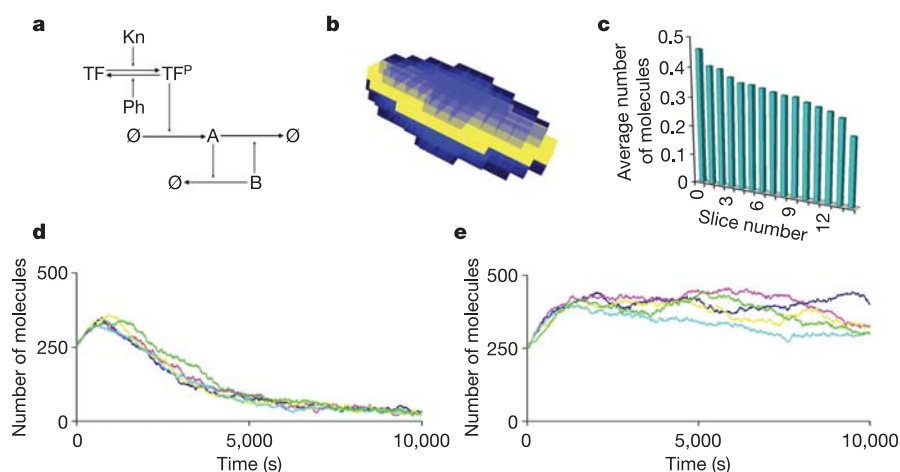


Figure 3 | Effect of localization of species on cellular processes. **a**, Diagram of a simple network in which a phosphorylated transcription factor TF^P triggers the synthesis of protein A that is involved in a positive feedback loop with protein B (resulting from mutual repression). $\emptyset \rightarrow$ indicates protein production, whereas $\rightarrow \emptyset$ indicates protein degradation. The behaviour will depend on the spatial constraints imposed on the kinase Kn and the phosphatase Ph. When they localize to opposite poles of the cell, switching

behaviour does not take place. Production of protein A in both simulations³⁷ takes place in the pole of the cell where the phosphatase is localized. **b**, The model geometry (lattice unit is $1 \mu m$). Yellow indicates the longitudinal cell slice along which the gradient is observed. **c**, Gradient of TF^P obtained with localization of species. **d**, **e**, Time course of protein B without (**d**) and with (**e**) localization of species (see Supplementary Fig. 4 for parameters used in the simulation).

overcoming data limitation by dividing the system into functional modules. One of the most beautiful paradigms of simulation–experiment interplay is in *E. coli* chemotaxis^{54,55}, but also significant discoveries have been made thanks to mathematical modelling in mammalian epidermal growth factor receptor (EGFR) and extracellular-signal-regulated kinase (ERK) signalling pathways^{56,57}.

As these recent examples show, the interplay between experiments and modelling has led to increasingly accurate models that not only fit the experimental data, but also make correct predictions and suggest new testable hypotheses.

Gaining understanding by mimicking nature

Natural systems can be discouragingly complex when tackled from a reverse-engineering approach. An alternative approach that makes extensive use of modelling and simulation tools has recently shown great promise: that of designing and building synthetic networks. Similar to the field of protein folding, where proteins are being successfully redesigned without a full understanding of how the proteins fold, biological circuits are being designed to engineer new

properties into living organisms, despite the limitations in our understanding of the cell processes.

Designs found in nature can be tested in isolation for a given function (Fig. 5). Well-characterized components help lower the barriers to modelling. The use of control elements (such as temperature for a temperature-sensitive protein, or an exogenous small molecule affecting a reaction) helps model validation by allowing a multitude of conditions to be tested. Also, the possibility to connect synthetic networks (for example, connecting a switch with a rheostat to make an oscillator) paves the way for investigating principles of modularity.

Synthetic networks are often patchworks of previously used parts, not only because a single point mutation may alter the *in vivo* activity of the network, but also because we cannot currently predict how redesigned molecules such as synthetic promoters will behave. Indeed, the better characterized the pieces are, the more confident one can be in the corresponding models (although spurious interactions can always arise owing to the complexity of cellular environments causing discrepancies between models and experiments). In fact, a central repository for well-characterized parts and modules and related information already exists (http://parts.mit.edu/registry/index.php/Main_Page), and is commendably helping to consolidate and expand knowledge in this field.

Designing a synthetic biological network involves a conscious choice of inputs and outputs—the handles and readouts of the system. As an example of a handle, compounds can be added that modulate the behaviour of specific components (for example, the Tet repressor can be inactivated by anhydrotetracycline), allowing the exploration of many configurations, including unnatural ones;

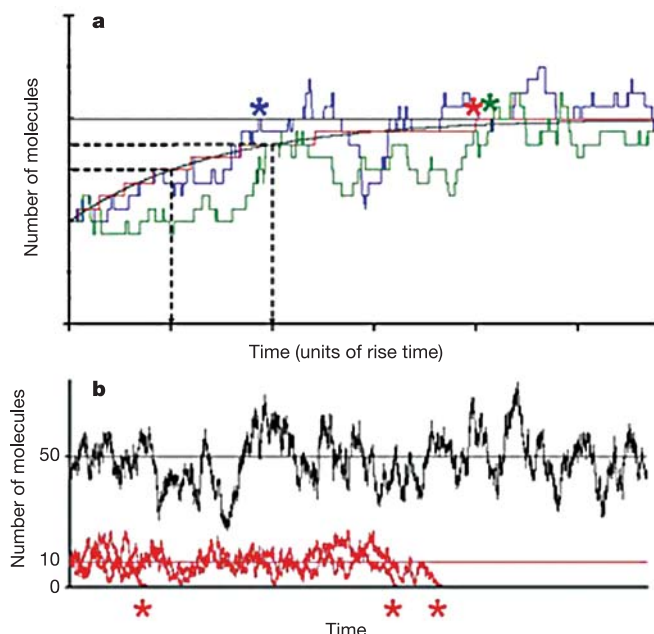


Figure 4 | Example of putative model-dependent pitfalls in modelling: continuous versus discrete concentrations. **a**, Simulation of a simple system in which a species is produced at a constant rate (16×10^{-3} molecules s^{-1}), and degraded with a first order decay rate of $1 \times 10^{-3} s^{-1}$ giving an asymptotic steady-state of 16 molecules per cell (plotted as a straight horizontal line). The initial state for all simulations is 8 molecules per cell. Two discrete stochastic runs are shown in blue and green. The corresponding ODE solution is shown in red (rounding each value of the continuous run to the closest integer). A star indicates, for each discrete run, when the steady-state value has been reached. Each broken line illustrates a geometric construction of the rise-time, as the time taken by the system to go, first from 8 to 4 molecules per cell away from steady state, then from 4 to 2. Each unit of rise-time is separated by tick-marks on the x-axis, which are identically spaced (as a property of exponential curves). The stochastic simulation shows that it is possible to reach steady-state values earlier or later than in the deterministic analysis. **b**, Superimposed stochastic (wavy lines) and deterministic (continuous straight lines) simulation runs of a system composed of a single autocatalytic species, with parameter values (see Supplementary Fig. 5) such that the upper steady-state solution is either high (black curves) or low (red curves) (respectively, 50 and 10 molecules per cell volume of $1 \mu m^3$). A star indicates when a stochastic run has reached the zero absorbing state. In the lower case (red curves), the deterministic solution clearly differs from the average stochastic run, because all runs reach zero and stay there.

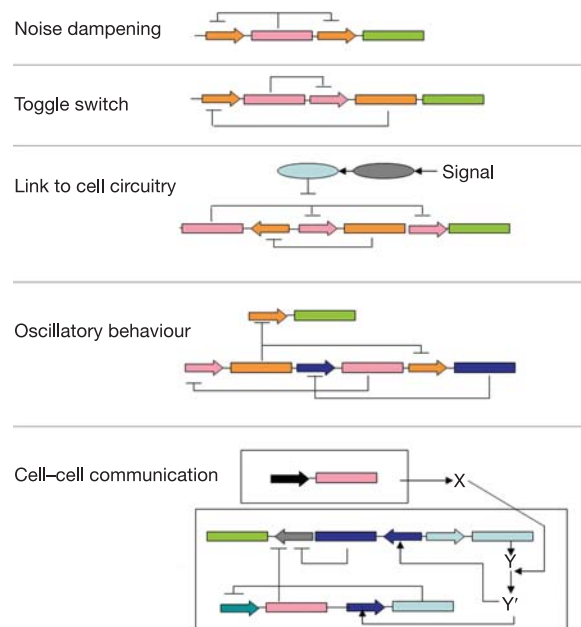


Figure 5 | Example of synthetic gene networks built with defined behaviour. Promoters are represented by arrows and genes by elongated boxes. Products of genes that are activated by another product are represented by X or Y. Cells are represented by large transparent boxes. Natural cell circuits are represented by ellipses. Reporter products are normally shown in green. When more than one bibliographic reference is given, only the circuitry used in the first cited reference is shown. Shown are: a noise-dampening negative feedback loop¹⁰; switch behaviour (toggle switch mutual repression^{65–67}, positive feedback loop^{68,69} and logic gates¹¹); interfacing natural and synthetic networks⁷⁰; oscillatory behaviour (repressilator⁷¹, metabolic oscillator⁷² and oscillator resistant to fluctuations⁷³); and cell–cell communication (regulation of protein expression⁷⁴, biofilm formation⁷⁰, population dynamics⁷⁵, selective killing of cells⁷⁵, pattern formation⁷⁶ and detection of gene transfer⁷⁷).

although the list of such known regulator/molecular target pairs is still limited, it is being expanded by dedicated efforts (for example, drug design). At the other end, the experimental readout is expected to give a static or dynamic picture of the state of the system, for a cell or a population, usually measuring the concentrations of key components. Appropriate detection methods are those that offer a good compromise between non-invasiveness, sensitivity, versatility, detection speed, ease of measurement, and so on, with champions of the trade such as green fluorescent protein, promising new candidates such as fluorescein arsenical hairpin binder (FLAsH)⁵⁸, and also a wealth of antibody and radioisotope-based methods to choose from. Finally, as most detection methods are indirect, carefully chosen internal standards are needed for comparing experimental data to model predictions (for example, fluorescence counts and number of molecules per cell).

For years, models of gene networks have been used to investigate general issues of relevance to molecular biology, but were awaiting proper experimental validation. The field of synthetic biology has clearly started addressing these issues with circuits of various designs⁵⁹ (Fig. 5) guided by their own collection of models, and many experimental results are corroborating simulation predictions.

From simulating to understanding: design principles

Contemporary experimental techniques offer insight about biological functions at the molecular level. Given the complexity of natural systems, however, knowledge of all interactions happening at this level does not generally provide the modeller with an intuitive and coherent comprehension about the process of interest. For instance, although we can understand how a signal is propagated in a linear cascade from cell membrane to nucleus, we find it difficult to make sense of a highly interconnected gene network, with the added complication that only part of the network may be active at any one time. One strategy is to decompose the network into more manageable modules, a task that is not trivial and somewhat arbitrary⁶⁰ because it requires grouping parts on the basis of the purported biological function achieved by the module within the network. Groupings are primarily suggested by topology, but as mentioned above, cases abound in which topology alone is insufficient to elicit function. Even close-to-ideal, single-input, single-output modules with a clear function, such as some signal transduction cascades seem to be, can require further specifications to predict important details of their operation, such as the topology-inferred property of ultrasensitivity, which was recently shown to depend on a set of conditions usually not met in real cascades⁶¹.

The search for and characterization of modular units within networks is an active field of systems biology where the observation of recurring designs linked to specific properties is providing another level of description of biological systems. A behaviour predicted for an isolated system, such as a small network motif (recurring topological pattern found over-represented in biological networks⁶²), is not straightforward to extrapolate to systems in their natural environment, given how intricately embedded they can be; consequently, despite the established modularity of networks, successful extrapolations⁶³ almost come as a surprise. These higher-level descriptions of how biological network modules behave draw heavily on analogies with the way human-engineered systems work. Engineering concepts about network components (for example, switches, amplifiers and control elements) and properties (for example, robustness and modularity) have provided invaluable rational descriptions of module behaviour by abstracting the details of molecular interactions.

Pushing the engineering analogy even further, systems biology studies have the aim of uncovering the 'organizational principles', or 'design principles', of biological systems. Although there are obvious radical differences between human-engineered and natural systems (for instance, engineers lay down a wish-list of intended properties for their system during the design phase), natural systems do have solutions similar to human-engineered ones in terms of certain

emergent properties (for example, modularity and noise attenuation), details of design (for example, feedback loops) and behaviour (for example, oscillations), as if conforming to a strict set of constraints that biologists are on a quest to discover (not to be confused with any universal law of nature, embedded, as if by magic, in each and every life form). Actually, beyond their natural appeal, the use of systems-theoretical concepts is perhaps our only chance to logically formulate the way a complex biological process operates in a concise, synthetic, human-understandable manner.

Although much remains to be learned in this field, simulation predictions of natural or engineered biological networks are helping us to identify the logical links between system design and system behaviour. This includes understanding the rationale behind the preferential use in nature of certain molecular blueprints to elicit certain elementary functions, such as the use of negative feedback loops to achieve homeostasis¹³, sign-sensitive delay properties of feed-forward loops⁶³, or more elaborate designs such as robust entrainment and temperature compensation of circadian clocks (achieved through densely connected loops⁶⁴), and the remarkable property of cellular 'adaptation', so simply formulated, and yet still so mysterious⁵⁶.

Perspectives

By discovering design principles, identifying biological modules, and quantitatively understanding how they operate through experiments and simulations, we hope to elucidate biological function as well as predict the effect of internal perturbations (for example, genetic mutations) or external perturbations (for example, drugs) so that disease treatments are more precise and effective. Similarly, understanding biological modules and being able to engineer new ones will pave the way for re-engineering of organisms and cells for numerous applications (including medical, agricultural and ecological situations). Thus, although a global and perfect understanding of a living system is not expected in the near future, the combination of modelling and experimentation offers the possibility of making inroads towards that goal, as well as developing new exciting, useful applications.

1. Kant, I. *Critique of Pure Reason*, University of Virginia Library, Electronic Text Center, Topic I, Part II, 45 (<http://etext.lib.virginia.edu/toc/modeng/public/KanPure.html>).
2. Locke, J. C. W. Extension of a genetic network model by iterative experimentation and mathematical analysis. *Mol. Syst. Biol.* doi:10.1038/msb4100018 (28 June 2005).
3. Albert, M. A. *et al.* Experimental and *in silico* analyses of glycolytic flux control in bloodstream form *Trypanosoma brucei*. *J. Biol. Chem.* **280**, 28306–28315 (2005).
4. Lee, E., Salic, A., Kruger, R., Heinrich, R. & Kirschner, M. W. The roles of APC and Axin derived from experimental and theoretical analysis of the Wnt pathway. *PLoS Biol.* doi:10.1371/journal.pbio.0000010 (13 October 2003).
5. di Bernardo, D. *et al.* Chemogenomic profiling on a genome-wide scale using reverse-engineered gene networks. *Nature Biotechnol.* **23**, 377–383 (2005).
6. Segre, D., Vitkup, D. & Church, G. M. Analysis of optimality in natural and perturbed metabolic networks. *Proc. Natl Acad. Sci. USA* **99**, 15112–15117 (2002).
7. Chang, P. L. Clinical bioinformatics. *Chang Gung Med. J.* **28**, 201–211 (2005).
8. Kerckhoffs, R. C. *et al.* Electromechanics of the paced left ventricle simulated by a straightforward mathematical model: comparison with experiments. *Am. J. Physiol. Heart Circ. Physiol.* **5**, H1889–H1897 (2005).
9. Dens, E. J., Bernaerts, K., Standaert, A. R. & Van Impe, J. F. Cell division theory and individual-based modeling of microbial lag: part I. The theory of cell division. *Int. J. Food Microbiol.* **101**, 303–318 (2005).
10. Becskei, A. & Serrano, L. Engineering stability in gene networks by autoregulation. *Nature* **405**, 590–593 (2000).
11. Guet, C. C., Elowitz, M. B., Hsing, W. & Leibler, S. Combinatorial synthesis of genetic networks. *Science* **296**, 1466–1470 (2002).
12. Shmulevich, I., Kauffman, S. A. & Aldana, M. Eukaryotic cells are dynamically ordered or critical but not chaotic. *Proc. Natl Acad. Sci. USA* **102**, 13439–13444 (2005).
13. Hardy, S. & Robillard, P. N. Modeling and simulation of molecular biology systems using Petri nets: modeling goals of various approaches. *J. Bioinform. Comput. Biol.* **2**, 595–613 (2004).
14. Errampalli, D. D., Priami, C. & Quaglia, P. A formal language for computational systems biology. *OMICS* **8**, 370–380 (2004).

15. Batt, G. *et al.* Validation of qualitative models of genetic regulatory networks by model checking: analysis of the nutritional stress response in *Escherichia coli*. *Bioinformatics* **21** (Suppl 1), i19–i28 (2005).
16. Kuipers, B. in *Readings in Qualitative Reasoning about Physical Systems* (ed. deKleer, D. S. W. J.) 257–274 (Morgan Kaufmann, San Francisco, 1989).
17. Csete, M. E. & Doyle, J. C. Reverse engineering of biological complexity. *Science* **295**, 1664–1669 (2002).
18. Louis, M. & Becskei, A. Binary and graded responses in gene networks. *Sci. STKE* **143**, PE33 (2002).
19. Clarke, B. L. Complete set of steady states for the general stoichiometric dynamical system. *J. Chem. Phys.* **75**, 4970–4979 (1981).
20. Edwards, J. S., Ibarra, R. U. & Palsson, B. O. *In silico* predictions of *Escherichia coli* metabolic capabilities are consistent with experimental data. *Nature Biotechnol.* **19**, 125–130 (2001).
21. Stelling, J., Klamt, S., Bettenbrock, K., Schuster, S. & Gilles, E. D. Metabolic network structure determines key aspects of functionality and regulation. *Nature* **420**, 190–193 (2002).
22. Reed, J. L., Vo, T. D., Schilling, C. H. & Palsson, B. O. An expanded genome-scale model of *Escherichia coli* K-12 (iJR904 GSM/GPR). *Genome Biol.* **4**, R54 (2003).
23. Fell, D. A. & Small, J. R. Fat synthesis in adipose tissue. An examination of stoichiometric constraints. *Biochem. J.* **238**, 781–786 (1986).
24. Covert, M. W., Schilling, C. H. & Palsson, B. Regulation of gene expression in flux balance models of metabolism. *J. Theor. Biol.* **213**, 73–88 (2001).
25. Burgard, A. P. & Maranas, C. D. Optimization-based framework for inferring and testing hypothesized metabolic objective functions. *Biotechnol. Bioeng.* **82**, 670–677 (2003).
26. Fall, C. P., Marland, E. S., Wagner, J. M. & Tyson, J. J. (eds) *Computational Cell Biology*. 1st edn. (Springer, 2002).
27. Savageau, M. A. Biochemical systems theory: operational differences among variant representations and their significance. *J. Theor. Biol.* **151**, 509–530 (1991).
28. Liu, Q. & Jia, Y. Fluctuations-induced switch in the gene transcriptional regulatory system. *Phys. Rev. E* **70**, 041907 (2004).
29. Thattai, M. & van Oudenaarden, A. Stochastic gene expression in fluctuating environments. *Genetics* **167**, 523–530 (2004).
30. Hanggi, P. Stochastic resonance in biology. How noise can enhance detection of weak signals and help improve biological information processing. *ChemPhysChem* **3**, 285–290 (2002).
31. Gillespie, D. T. General method for numerically simulating the stochastic time evolution of coupled chemical reactions. *J. Comput. Phys.* **22**, 403–434 (1976).
32. Gillespie, D. T. Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.* **81**, 2340–2361 (1977).
33. Haseltine, E. L. & Rawlings, J. B. Approximate simulation of coupled fast and slow reactions for stochastic chemical kinetics. *J. Chem. Phys.* **117**, 6959–6969 (2002).
34. Rathinam, M., Petzold, L. R., Cao, Y. & Gillespie, D. T. Stiffness in stochastic chemically reacting systems: the implicit tau-leaping method. *J. Chem. Phys.* **119**, 12784–12794 (2003).
35. Rao, C. V. & Arkin, A. P. Stochastic chemical kinetics and the quasi-steady-state assumption: application to the Gillespie algorithm. *J. Chem. Phys.* **118**, 4999–5010 (2003).
36. Stundzia, A. B. & Lumsden, C. J. Stochastic simulation of coupled reaction-diffusion processes. *J. Comput. Phys.* **127**, 196–207 (1996).
37. Ander, M. *et al.* SmartCell, a framework to simulate cellular processes that combines stochastic approximation with diffusion and localisation: analysis of simple networks. *Systems Biol.* **1**, 129–138 (2004).
38. Bezrukov, S. M., Frauenfelder, H. & Moss, F. (eds) *Fluctuations and Noise in Biological, Biophysical, and Biomedical Systems* (Proc. SPIE, Vol. 5110, 2003).
39. Salis, H. & Kaznessis, Y. Accurate hybrid stochastic simulation of a system of coupled chemical or biochemical reactions. *J. Chem. Phys.* **122**, 054103 (2005).
40. Alfonsi, A., Cancas, E., Turinici, G., Di Ventura, B. & Huisinga, W. Adaptive simulation of hybrid stochastic and deterministic models for biochemical systems. *ESAIM Proc.* **14**, 1–13 doi:10.1051/proc:2005001 (2005).
41. Gitai, Z. The new bacterial cell biology: moving parts and subcellular architecture. *Cell* **120**, 577–586 (2005).
42. Gorlich, D., Seewald, M. J. & Ribbeck, K. Characterization of Ran-driven cargo transport and the RanGTPase system by kinetic measurements and computer simulation. *EMBO J.* **22**, 1088–1100 (2003).
43. Nedelec, F., Surrey, T. & Karsenti, E. Self-organisation and forces in the microtubule cytoskeleton. *Curr. Opin. Cell Biol.* **15**, 118–124 (2003).
44. Savai, S., Thomason, P. A. & Cox, E. C. An autoregulatory circuit for long-range self-organization in *Dictyostelium* cell populations. *Nature* **433**, 323–326 (2005).
45. Collier, J. R., Monk, N. A., Maini, P. K. & Lewis, J. H. Pattern formation by lateral inhibition with feedback: a mathematical model of Delta–Notch intercellular signalling. *J. Theor. Biol.* **183**, 429–446 (1996).
46. Wu, D., Jia, Y., Yang, L., Liu, Q. & Zhan, X. Phase synchronization and coherence resonance of stochastic calcium oscillations in coupled hepatocytes. *Biophys. Chem.* **115**, 37–47 (2005).
47. Lemerle, C., Di Ventura, B. & Serrano, L. Space as the final frontier in stochastic simulations of biological systems. *FEBS Lett.* **579**, 1789–1794 (2005).
48. Ellis, R. J. Macromolecular crowding: obvious but underappreciated. *Trends Biochem. Sci.* **26**, 597–604 (2001).
49. Yarmush, M. L. & Banta, S. Metabolic engineering: advances in modeling and intervention in health and disease. *Annu. Rev. Biomed. Eng.* **5**, 349–381 (2003).
50. Price, N. D., Reed, J. L. & Palsson, B. O. Genome-scale models of microbial cells: evaluating the consequences of constraints. *Nature Rev. Microbiol.* **2**, 886–897 (2004).
51. Fong, S. S. & Palsson, B. O. Metabolic gene-deletion strains of *Escherichia coli* evolve to computationally predicted growth phenotypes. *Nature Genet.* **36**, 1056–1058 (2004).
52. Alvarez-Vasquez, F. *et al.* Simulation and validation of modelled sphingolipid metabolism in *Saccharomyces cerevisiae*. *Nature* **433**, 425–430 (2005).
53. Klipp, E., Nordlander, B., Kruger, R., Gennemark, P. & Hohmann, S. Integrative model of the response of yeast to osmotic shock. *Nature Biotechnol.* **23**, 975–982 (2005).
54. Abouhamad, W. N. *et al.* Computer-aided resolution of an experimental paradox in bacterial chemotaxis. *J. Bacteriol.* **180**, 3757–3764 (1998).
55. Kalir, S. & Alon, U. Using a quantitative blueprint to reprogram the dynamics of the flagella gene network. *Cell* **117**, 713–720 (2004).
56. Wiley, H. S., Shvartsman, S. Y. & Lauffenburger, D. A. Computational modeling of the EGF-receptor system: a paradigm for systems biology. *Trends Cell Biol.* **13**, 43–50 (2003).
57. Sasagawa, S., Ozaki, Y., Fujita, K. & Kuroda, S. Prediction and validation of the distinct dynamics of transient and sustained ERK activation. *Nature Cell Biol.* **7**, 365–373 (2005).
58. Martin, B. R., Giepmans, B. N., Adams, S. R. & Tsien, R. Y. Mammalian cell-based optimization of the biarsenical-binding tetracycline motif for improved fluorescence and affinity. *Nature Biotechnol.* **23**, 1308–1314 (2005).
59. Hasty, J., McMillen, D. & Collins, J. J. Engineered gene circuits. *Nature* **420**, 224–230 (2002).
60. Levine, M. & Davidson, E. H. Gene regulatory networks for development. *Proc. Natl Acad. Sci. USA* **102**, 4936–4942 (2005).
61. Ortega, F., Acerenza, L., Westerhoff, H. V., Mas, F. & Cascante, M. Product dependence and bifunctionality compromise the ultrasensitivity of signal transduction cascades. *Proc. Natl Acad. Sci. USA* **99**, 1170–1175 (2002).
62. Milo, R. *et al.* Network motifs: simple building blocks of complex networks. *Science* **298**, 824–827 (2002).
63. Mangan, S., Zaslaver, A. & Alon, U. The coherent feedforward loop serves as a sign-sensitive delay element in transcription networks. *J. Mol. Biol.* **334**, 197–204 (2003).
64. Schoning, J. C. & Staiger, D. At the pulse of time: protein interactions determine the pace of circadian clocks. *FEBS Lett.* **579**, 3246–3252 (2005).
65. Atkinson, M. R., Savageau, M. A., Myers, J. T. & Ninfa, A. J. Development of genetic circuitry exhibiting toggle switch or oscillatory behaviour in *Escherichia coli*. *Cell* **113**, 597–607 (2003).
66. Gardner, T. S., Cantor, C. R. & Collins, J. J. Construction of a genetic toggle switch in *Escherichia coli*. *Nature* **403**, 339–342 (2000).
67. Kramer, B. P. *et al.* An engineered epigenetic transgene switch in mammalian cells. *Nature Biotechnol.* **22**, 867–870 (2004).
68. Becskei, A., Seraphin, B. & Serrano, L. Positive feedback in eukaryotic gene networks: cell differentiation by graded to binary response conversion. *EMBO J.* **20**, 2528–2535 (2001).
69. Isaacs, F. J., Hasty, J., Cantor, C. R. & Collins, J. J. Prediction and measurement of an autoregulatory genetic module. *Proc. Natl Acad. Sci. USA* **100**, 7714–7719 (2003).
70. Kobayashi, H. *et al.* Programmable cells: interfacing natural and engineered gene networks. *Proc. Natl Acad. Sci. USA* **101**, 8414–8419 (2004).
71. Elowitz, M. B. & Leibler, S. A synthetic oscillatory network of transcriptional regulators. *Nature* **403**, 335–338 (2000).
72. Fung, E. *et al.* A synthetic gene-metabolic oscillator. *Nature* **435**, 118–122 (2005).
73. Hasty, J., Dolnik, M., Rottschäfer, V. & Collins, J. J. Synthetic gene network for entraining and amplifying cellular oscillations. *Phys. Rev. Lett.* **88**, 148101 (2002).
74. Bulter, T. *et al.* Design of artificial cell–cell communication using gene and metabolic networks. *Proc. Natl Acad. Sci. USA* **101**, 2299–2304 (2004).
75. You, L., Cox, R. S. III, Weiss, R. & Arnold, F. H. Programmed population control by cell–cell communication and regulated killing. *Nature* **428**, 868–871 (2004).
76. Basu, S., Gerchman, Y., Collins, C. H., Arnold, F. H. & Weiss, R. A synthetic multicellular system for programmed pattern formation. *Nature* **434**, 1130–1134 (2005).
77. Jaenecke, S., de Lorenzo, V., Timmis, K. N. & Diaz, E. A stringently controlled expression system for analysing lateral gene transfer between bacteria. *Mol. Microbiol.* **21**, 293–300 (1996).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank J. Reich, J. J. Collins, I. Mattaj and J. Gagneur for critical reading of the manuscript, and P. Beltrao and M. Isalan for helpful discussions. This work was financed partly by the EC grant, COMBIO.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence should be addressed to L.S. (serrano@embl.de).

ARTICLES

Transiting extrasolar planetary candidates in the Galactic bulge

Kailash C. Sahu¹, Stefano Casertano¹, Howard E. Bond¹, Jeff Valenti¹, T. Ed Smith¹, Dante Minniti², Manuela Zoccali², Mario Livio¹, Nino Panagia¹, Nikolai Piskunov³, Thomas M. Brown¹, Timothy Brown⁴, Alvio Renzini⁵, R. Michael Rich⁶, Will Clarkson¹ & Stephen Lubow¹

More than 200 extrasolar planets have been discovered around relatively nearby stars, primarily through the Doppler line shifts owing to reflex motions of their host stars, and more recently through transits of some planets across the faces of the host stars. The detection of planets with the shortest known periods, 1.2–2.5 days, has mainly resulted from transit surveys which have generally targeted stars more massive than $0.75 M_{\odot}$, where $M_{\odot}$ is the mass of the Sun. Here we report the results from a planetary transit search performed in a rich stellar field towards the Galactic bulge. We discovered 16 candidates with orbital periods between 0.4 and 4.2 days, five of which orbit stars of masses in the range 0.44 – $0.75 M_{\odot}$. In two cases, radial-velocity measurements support the planetary nature of the companions. Five candidates have orbital periods below 1.0 day, constituting a new class of ultra-short-period planets, which occur only around stars of less than $0.88 M_{\odot}$. This indicates that those orbiting very close to more-luminous stars might be evaporatively destroyed or that jovian planets around stars of lower mass might migrate to smaller radii.

Radial-velocity (RV) searches^{1–3} have led to the unexpected discovery of jovian-mass exoplanets with orbital periods of only a few days—‘hot Jupiters’. RV surveys favour stars bright enough for high-resolution spectroscopy and have thus generally been limited to nearby stars that are hotter than early K spectral type, although RV studies are now being extended to M dwarfs^{4,5}. Photometric monitoring of microlensing events provides another technique that has recently led to the discovery of a planet of 5.5 Earth masses around a low-mass star⁶.

The increasingly successful transit surveys are based on detecting the recurrent dimming of the primary star caused by the silhouetted planet transiting across its face. The first transiting planet, HD 209458b, was found through photometric follow-up of a RV detection^{7,8}. Subsequently, five of the transiting candidates found by the Optical Gravitational Lensing Experiment (OGLE) survey⁹ have been confirmed as having planetary masses through RV measurements^{10–14}, and four other planets around bright field stars have been found in separate transit surveys^{15–18}. These ten RV-confirmed transiting planets have periods between 1.2 and 4.0 days, masses of 0.36 – $1.45 M_J$ (where M_J is the mass of Jupiter), and radii of 0.7 – $1.4 R_J$ (where R_J is the radius of Jupiter).

Transits occur only when the planetary orbit is viewed nearly edge-on. Because of the low probability of such an orientation¹⁹, transit surveys must necessarily cover large numbers of stars, either through wide-angle imaging or through deep imaging in smaller fields with high stellar densities. Because planetary transit depths are rarely more than a few per cent, high-precision photometry is required in transit searches. So far, the deepest ground-based transit detections have reached $V \approx 17$, limiting planet discovery to stars more massive than $\sim 0.75 M_{\odot}$, lying within ~ 2 kpc of the Sun.

The Sagittarius Window Eclipsing Extrasolar Planet Search (SWEPS) project reported here was conceived as a transit survey that would fully exploit the high spatial resolution and high photometric precision capabilities of the Hubble Space Telescope (HST).

We used HST’s Advanced Camera for Surveys (ACS) and chose a rich field, lying in the Sagittarius window of the Galactic bulge, which we monitored for periodic small dimmings over a continuous seven-day period.

We show that at least 7 of our 16 candidates are likely to be genuine planets, rather than brown dwarfs or low-mass stars. Because of the faintness and crowding of the targets, RV confirmation is currently not feasible for the great majority of the targets. However, for two of our brightest candidates we have used RVs to place limits on the mass of the transiting objects, ruling out stellar companions.

Stellar properties in the SWEPS field

The SWEPS field lies in the Sagittarius I Window of the Galactic bulge, where the stellar population has a wide variety of metallicities, ranging over $-1.5 < [\text{Fe}/\text{H}] < +0.5$ (refs 20, 21). Our HST search can detect planetary transits of stars as faint as $V \sim 26$, which at the distance of the Galactic bulge corresponds to a main-sequence mass of $\sim 0.45 M_{\odot}$.

We monitored the SWEPS field over the seven-day interval 2004 February 22–29, in the V (F606W) and I (F814W) filters. Combination of all of the exposures produces extremely deep V and I images, in which 245,000 stars are detected to $V \sim 30$ (see Methods for calibration details). The colour–magnitude diagram (CMD; Fig. 1) shows two components: a dominant population of old stars with a main-sequence turnoff near $V = 20$ exhibiting well-populated sub-giant and giant branches, and a much less numerous, closer, and younger population with an unevolved main sequence. We associate the old population with the Galactic bulge, and the younger objects with the foreground disk population, the latter being dominated by the Sagittarius spiral arm^{22,23}.

The magnitude and colour of each star can be used to estimate its mass, temperature and radius^{24–26}. The bulge component is well represented by a population with a distance modulus of $(m - M)_0 = 14.3$, a reddening of $E(B - V) = 0.64$, and a mean

¹Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, Maryland 21218, USA. ²Universidad Catolica de Chile, Av. Vicuña Mackenna 4860, Santiago 22, Chile.

³Uppsala University, Box 515, S-751 20 Uppsala, Sweden. ⁴High Altitude Observatory, 3450 Mitchell Lane, Boulder, Colorado 80307, USA. ⁵INAF-Osservatorio Astronomico di Padova, Vicolo dell’Osservatorio 5, 35122 Padova, Italy. ⁶University of California at Los Angeles, Los Angeles, California 90095-1562, USA.

chemical composition of $[\text{Fe}/\text{H}] = 0.0$ and an α -element enhancement of $[\alpha/\text{Fe}] = 0.3$, similar to values found earlier^{27,28}. We adopt a nominal age of 10 Gyr, the isochrone for which is shown as a solid red line in Fig. 1. The blue dashed line shows a representative foreground disk population with an unevolved main sequence at a distance modulus of $(m - M)_0 = 13.1$ and the same reddening as the bulge. Both populations have a significant dispersion in magnitude, due primarily to their depth in the line of sight, and also to the metallicity dispersion in the bulge.

Detection and screening of candidates

In the SWEEPS field there are 180,000 dwarf stars brighter than $V = 27$. We determined a photometric time sequence for each of them by subtracting the total combined image from each individual frame and then measuring the residual flux at the location of each star in each difference image. We used the box-fitting least-squares (BLS) algorithm²⁹ to search for transits by calculating a BLS frequency spectrum for each star's time-series data. This algorithm provides approximate transit parameters (period, duration, zero-phase and transit depth), from which we estimated the signal-to-noise (S/N) ratio of the possible transit signal and retained only those with $S/N > 5$. For these initial candidates, we then refined the fitting parameters through a χ^2 minimization procedure, now using a code³⁰ that calculates a full synthetic model light curve. We retained as genuine candidates only those with a statistical significance sufficiently high that the total number of expected statistical false positives in the entire sample was less than one (see Methods).

The resulting list of systems with transiting companions was then screened to remove binaries in which the companions were likely to be low-mass stars rather than planets. We rejected objects showing

any of the following properties: a transit depth implying a companion radius of more than $1.4R_J$ (the radii of known transiting exoplanets range from $0.7R_J$ to $1.4R_J$; larger objects, even at short orbital separations, are likely to be stars^{31,32}); ellipsoidal light variations, implying tidal distortion of the primary by a companion of stellar mass; and secondary eclipses and/or different transit depths in V and I , indicating a significant light contribution from the companion. We also eliminated objects in which the photo-centre of the transit signal (in the difference image) was offset with respect to that of the uneclipsed star (in the direct image), implying the presence of an eclipsing stellar binary system whose light was blended with that of a very nearby third star. As an additional check for stellar eclipses, we doubled the period and recalculated the transit depths. Candidates with statistically significant differences in transit depths at phase 0.5 and 1.0 in this doubled period were rejected as probable stellar binaries.

This process of elimination resulted in a final list of 16 candidates. The magnitudes of their host stars range from $V = 18.8$ to $V = 26.2$, corresponding to stellar masses of 1.24 – $0.44 M_\odot$. Figure 2 shows a few examples of the observed transit light curves (see Supplementary Figs S1 and S2 for all of the light curves).

The rejected candidates include 165 eclipsing binaries in which the secondary is likely to be a stellar companion (K.C.S. *et al.*, unpublished observations), out of which 125 show ellipsoidal variations.

Are the transiting bodies really planets?

There are, however, several other astrophysical situations that can potentially produce shallow light-curve dips mimicking planetary transits³³, which we now discuss.

Grazing stellar eclipses. A stellar binary with a grazing eclipse can

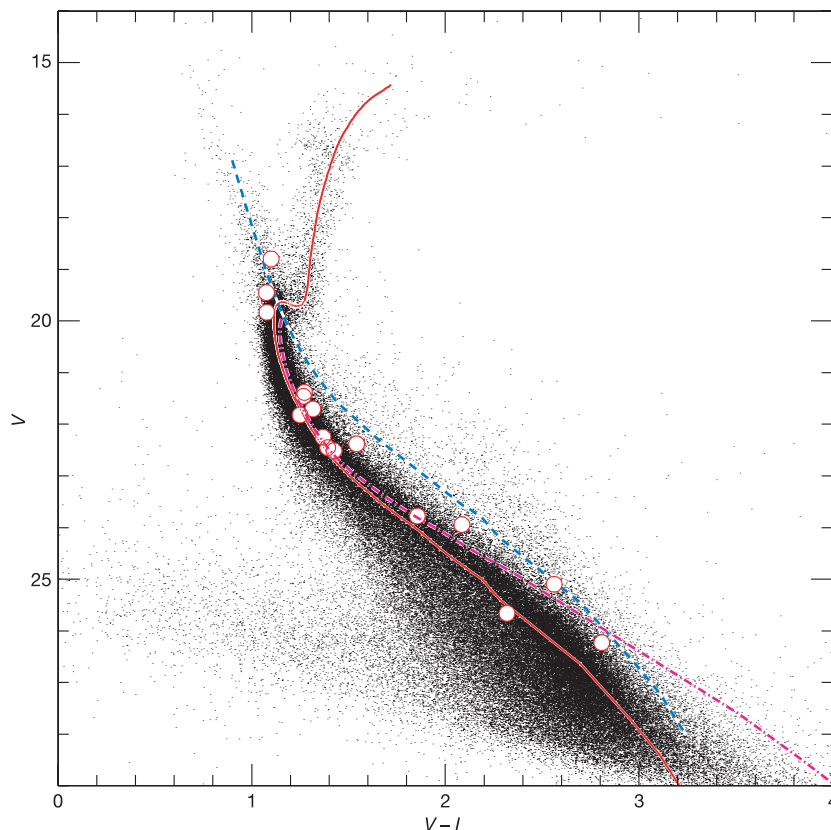


Figure 1 | CMD of the SWEEPS $202'' \times 202''$ field in the Galactic bulge. The CMD of the 245,000 stars was derived from the deep, combined ACS images, with total integration times of 86,106 and 89,835 s in the F606W (V) and F814W (I) filters, respectively. A solar-metallicity isochrone with an age of 10 Gyr (solid red line) fits the dominant bulge population; an unevolved

main sequence (dashed blue line) representative of the foreground young disk population is also shown. An estimated isochrone for $[\text{Fe}/\text{H}] = 0.5$ is shown by the magenta dot-dashed curve. Large circles mark the 16 host stars with transiting planetary candidates.

produce a depth similar to that due to a planetary transit. The expected number of grazing incidences can be predicted from the properties of the 40 stellar eclipsing binaries that do not show ellipsoidal variations. By assuming that these are drawn from a population with randomly distributed inclinations, and taking the binary parameters and the detection efficiencies associated with these systems into account, we estimate that a maximum of 1.4 of the 16 candidates could be a grazing system masquerading as a planetary transit.

Blended eclipsing binaries. A deeply eclipsing stellar binary, whose light is blended with a brighter constant star, can produce an eclipse in the combined light with a depth similar to that due to a planetary companion of a single star. Blending can occur either because of a chance overlap along the line of sight that causes no detectable shift of the photo-centre of the transit signal or because of membership in a physical triple.

The number of chance overlaps can be estimated from the 165 detected isolated eclipsing stellar binaries down to $V = 27$. On the basis of their surface density, we expect only 0.3 overlaps with brighter field stars down to $V = 24$, in contrast to the 13 detected planetary candidates that are associated with stars this bright. Assuming that the eclipsing-binary fraction is the same down to $V = 29$, we now expect ~ 0.8 candidates to be due to chance overlaps in the population down to $V = 27$.

About 10% of all binaries are in physical triple systems, according to statistics in the solar neighbourhood³⁴. Thus, on the basis of our 165 detected eclipsing binaries, there will be about 17 that are in

triples. However, for dilution to produce a shallow transit resembling that of a planet, the third star must lie in a narrow range of magnitudes brighter than the eclipsing pair. If we assume that the third star is drawn from the luminosity function observed in the SWEEPS field, only $\sim 10\%$ of these triples will have transits with depths resembling those due to planets; thus, of our 16 planetary candidates, ~ 2 could be due to physical triple systems.

Radius ambiguity. An ambiguity arises from the fact that, for masses between $\sim 0.5M_J$ and $\sim 150M_J$, planets, brown dwarfs and low-mass stars all have similar radii^{31,34,35}. Thus the transit depths that we have measured are insufficient by themselves to distinguish between the three types of companions. However, below we provide some statistical arguments indicating that a large majority of our transiting companions are in fact likely to be of planetary mass.

To exclude brown dwarfs as significant contaminants of our candidate sample, we note that RV surveys of nearby bright stars show a paucity of brown-dwarf companions in the range $13\text{--}80M_J$ (ref. 36)—the so-called ‘brown-dwarf desert’. If the same statistics hold for our bulge stars, the number of brown-dwarf companions will be negligible.

To assess the expected number of low-mass stellar companions in our candidate sample, we refer to results from the RV follow-up of the OGLE transit candidates. In the OGLE Galactic-bulge and Carina fields, a complete sample of about 60 transiting candidates has been followed up with RV measurements^{11–13,37,38}. Most of the screening process we were able to perform for the SWEEPS transits is not

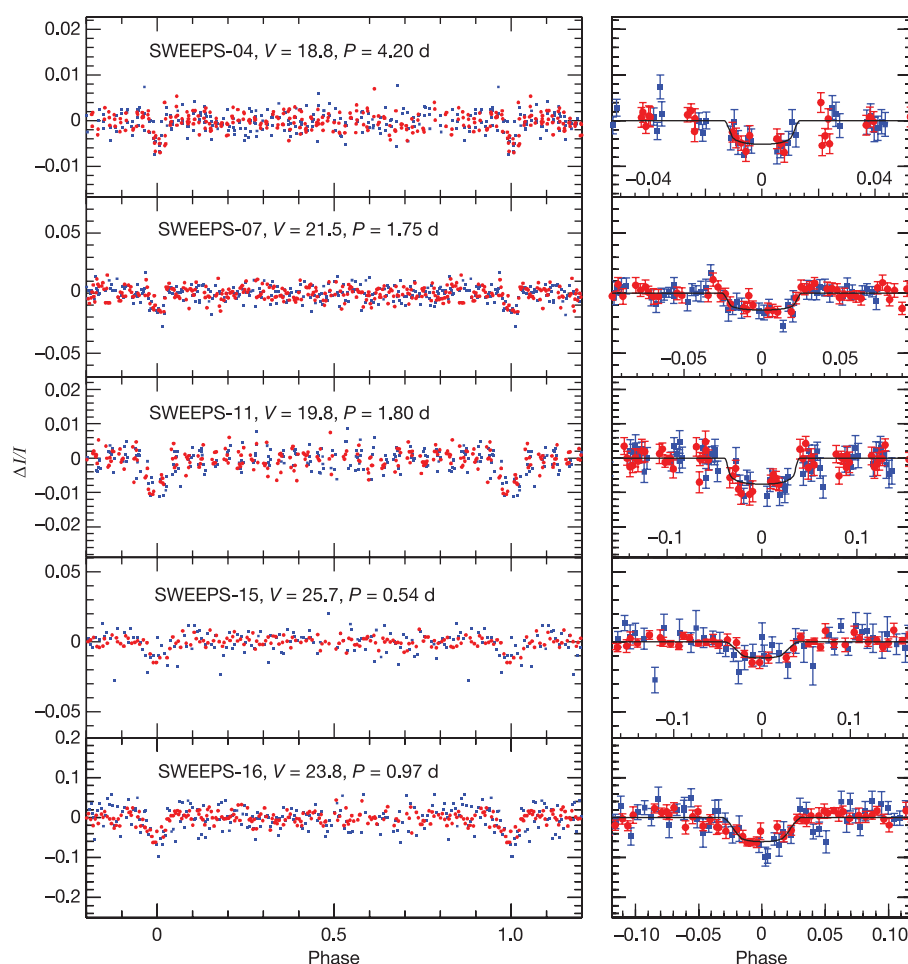


Figure 2 | Example transit light curves. Five examples of observed light curves, including two with RV measurements and two USPPs. The left panels show the entire light curve, phased at the derived orbital period, and the right panels show magnified views of the transit with $\pm 1\sigma$ error bars on the individual points. No binning is applied to the data, and the scatter is

generally consistent with photon noise. Blue squares are the V-band observations, and red circles are the I-band observations. The solid black curves are the best-fitting model transit light curves. Note that the fainter sources are intrinsically red, so they have a significantly higher S/N in the I band.

feasible for the OGLE candidate stars, because of their uncertain distances and stellar properties, the lack of multicolour photometry and the low spatial resolution of ground-based imaging. Nevertheless, the OGLE RV follow-up has yielded five confirmed transiting planetary companions, all of which have radii smaller than the SWEEPS cutoff of $1.4R_J$. All of the confirmed OGLE planets have orbital periods ≤ 4.0 days, and thus occupy the period range to which the SWEEPS survey is sensitive. In this period range, the RV follow-up measurements yielded one confirmed low-mass stellar companion, OGLE-TR-122b, which has a radius of $1.3R_J$ (ref. 35). One more low-mass stellar companion (OGLE-TR-123b) with a radius of $1.2R_J$ and a mass of $96M_J$ was found in the OGLE follow-up at a longer period of 7.2 days (ref. 38). From the existing data, we cannot exclude the possibility that there might be a higher proportion of small stellar companions around low-mass dwarfs than around G–K dwarfs. However, assuming that similar statistics apply to the Galactic bulge population, we infer that the stellar contamination rate in the SWEEPS sample is likely to be no more than $\sim 29\%$ (two out of seven), and may be half of that because the period of one of the contaminating low-mass OGLE stars with a planetary-sized radius is larger than that of the planets.

Cumulative sample contamination. Taking all of the possible contaminants into account, we estimate that at least 45% of our transiting candidates are genuine planets. Considering that no such contamination was found in the 47 Tuc data (see below), this is likely to be a conservative limit. The physical properties of the 16 planetary candidates are listed in Table 1. The Table shows that the radii of the SWEEPS planetary candidates range from a sub-saturnian $0.8R_J$ up to our imposed cutoff at $1.4R_J$.

Comparison with globular cluster 47 Tucanae

Strikingly different results were obtained in an intensive HST search for transiting planets performed in 1999 in the globular cluster 47 Tucanae³⁹. About 34,000 cluster stars were monitored throughout an eight-day period. In spite of similar stellar crowding, transit-depth sensitivity and mass range of the surveyed stars, the population of transiting planetary candidates found in the Galactic bulge is conspicuously absent from the cluster. We emphasize that 11 eclipsing stellar binaries were found in 47 Tuc (ref. 40) but not a single transiting object with a radius consistent with a planetary companion.

The lack of planets in 47 Tuc could be due either to its lower metallicity (Fe/H $\sim 20\%$ of solar) or to its high stellar density. The strong correlation between the planet frequency and stellar metallicity seen in the solar neighbourhood^{41,42} points towards metallicity as the cause, and provides a plausible explanation for the pronounced

difference in the planetary incidence between 47 Tuc and the Galactic bulge.

Mass limits from RV measurements

Although we argue that most of our transiting candidates are of planetary mass, unambiguous confirmation can come only from RV measurements. These are very challenging with current technology in the SWEEPS field because of the faintness and crowding. We have nonetheless been able to place meaningful limits on the companion masses of two of our brightest candidates, SWEEPS-04 and SWEEPS-11, which lie in relatively uncrowded locations. During 2004 June 22–25, we observed both stars simultaneously by using a fibre-fed mode of the UVES echelle spectrograph at the 8-m Very Large Telescope of the European Southern Observatory at Cerro Paranal, Chile. Two detectors covered the wavelength ranges 4,812–5,750 and 5,887–6,759 Å. In the total exposure time of 31 h we attained a peak S/N ratio of 7–10 per resolution element in the coadded spectrum of each star.

We found 51 stellar absorption features strong enough to provide a useful constraint on RV. We fitted each individual feature with a shifted and binned solar spectrum that had been numerically degraded to match our velocity resolution of 5 km s^{-1} . We then rejected outliers and calculated a noise-weighted mean RV for 11 separate epochs.

Figure 3 shows our measured RVs and formal uncertainties for each epoch. For SWEEPS-04, the absence of detected stellar velocity perturbations rules out at the 95% confidence level a companion more massive than $3.8M_J$ ($K > 0.42 \text{ km s}^{-1}$). For SWEEPS-11, we probably detect a RV variation with an amplitude of $1.5^{+0.9}_{-0.7} \text{ km s}^{-1}$, corresponding to a companion mass of $9.7^{+5.6}_{-4.5} M_J$, but a non-detection is also possible if residual systematic errors are significant. Setting aside the low probability of triple systems, the RV data indicate that both of these companions are planets (or a low-mass brown dwarf for SWEEPS-11), but certainly not stars.

Ultra-short-period planets

Five of our candidates have periods of less than 1.0 day. We call them ultra-short-period planets (USPPs), noting that the shortest orbital period yet found among RV-confirmed planets is 1.2 days, whereas our USPPs extend the periods down to 0.42 day. Statistical analysis of possible false positives indicates that at least two of these USPPs might be genuine planets.

All five USPPs orbit stars of less than $0.88 M_\odot$. USPPs thus seem to be analogues of hot Jupiters, but around lower-mass stars. We note that USPPs are not expected to be especially hot

Table 1 | The transiting planetary candidates

ID	RA (2000)	Dec. (2000)	Transit depth	Period (d)	S/N	Stellar mass ($M_\odot$)	Stellar radius ($R_\odot$)	V	I	Planet radius (R_J)	Error† in R_p (R_J)	Orbital radius (AU)	Orbital radius (R_*)
SWEEPS-01	17 h 58 min 53.29 s	−29° 12′ 33.5″	0.019	1.566	8.5	0.81	0.75	22.25	20.88	1.01	0.13	0.025	7.08
SWEEPS-02	17 h 58 min 53.38 s	−29° 12′ 17.8″	0.079	0.912	13.2	0.55	0.50	25.10	22.53	1.37	0.25	0.015	6.48
SWEEPS-03	17 h 58 min 53.57 s	−29° 11′ 44.1″	0.015	1.279	8.3	0.79	0.72	22.51	21.09	0.87	0.11	0.021	6.35
SWEEPS-04	17 h 58 min 53.92 s	−29° 11′ 20.6″	0.005	4.200	8.7	1.24	1.18	18.80	17.70	0.81	0.20	0.055	9.93
SWEEPS-05	17 h 58 min 54.60 s	−29° 11′ 28.2″	0.034	2.313	8.0	0.66	0.61	23.94	21.85	1.09	0.10	0.030	10.57
SWEEPS-06	17 h 58 min 57.29 s	−29° 12′ 53.4″	0.004	3.039	9.5	1.09	1.36	19.45	18.37	0.82	0.21	0.042	6.68
SWEEPS-07	17 h 58 min 57.69 s	−29° 11′ 14.5″	0.012	1.747	10.4	0.90	0.85	21.46	20.19	0.90	0.11	0.027	6.93
SWEEPS-08	17 h 58 min 59.24 s	−29° 13′ 28.7″	0.015	0.868	14.1	0.87	0.81	21.70	20.39	0.98	0.09	0.017	4.50
SWEEPS-09	17 h 58 min 59.60 s	−29° 12′ 11.8″	0.020	1.617	9.5	0.79	0.73	22.45	21.06	1.01	0.12	0.025	7.38
SWEEPS-10	17 h 59 min 02.00 s	−29° 13′ 23.7″	0.096	0.424	11.6	0.44	0.41	26.23	23.42	1.24	0.23	0.008	4.41
SWEEPS-11	17 h 59 min 02.67 s	−29° 11′ 53.5″	0.006	1.796	15.2	1.10	1.45	19.83	18.75	1.13	0.21	0.030	4.41
SWEEPS-12	17 h 59 min 04.44 s	−29° 13′ 17.1″	0.014	2.952	8.7	0.86	0.80	21.82	20.57	0.91	0.11	0.038	10.33
SWEEPS-13	17 h 59 min 05.95 s	−29° 13′ 05.6″	0.009	1.684	7.5	0.91	0.86	21.38	20.11	0.78	0.12	0.027	6.67
SWEEPS-14	17 h 59 min 07.56 s	−29° 10′ 39.8″	0.017	2.965	8.0	0.80	0.73	22.38	20.84	0.93	0.09	0.037	10.98
SWEEPS-15	17 h 59 min 07.64 s	−29° 10′ 23.7″	0.099	0.541	8.5	0.49	0.45	25.66	23.34	1.37	0.30	0.010	4.90
SWEEPS-16	17 h 59 min 08.44 s	−29° 11′ 40.6″	0.053	0.969	15.3	0.68	0.62	23.78	21.92	1.40	0.18	0.017	5.83

† The error in the planet radius (R_p) takes into account the error in the stellar radius caused by the dispersion in the CMD and the error due to transit light-curve fitting, but does not take into account other possible systematics.

compared with previously known ‘hot Jupiters’, because the irradiance from the low-mass primary at their locations is comparable to that of planets found around more massive stars. In fact, we argue below that irradiance levels might be one of the reasons why USPPs are not found around more massive stars. We also note that, in units of host stellar radii, the USPPs are no closer to their parents than the closest of the ordinary hot Jupiters. For example, the smallest orbital separation in units of stellar radii (R_*) among the SWEEPS candidates is that of the USPP SWEEPS-10 ($4.412R_*$), whereas the next smallest is that of SWEEPS-11, a 1.8-day hot Jupiter orbiting a $1.1 M_\odot$ star at $4.414R_*$. The shortest-period RV-confirmed planet, OGLE-TR-56b, has an orbital radius of $4.397R_*$.

USPPs do not raise an issue of stability against tidal breakup, because even at the closest of the observed separations a planet of more than $1.6M_J$ will lie within its Roche radius.

Physics of the planetary candidates

The positions of the 16 host stars in the CMD of the SWEEPS field are marked with open circles in Fig. 1. Below $V \sim 23$ there is an increasing tendency for the hosts to lie redward of the bulge ridge-line. As shown in Fig. 1, the high-metallicity isochrone (dash-dotted magenta curve) also exhibits a progressively increasing shift redward as one moves

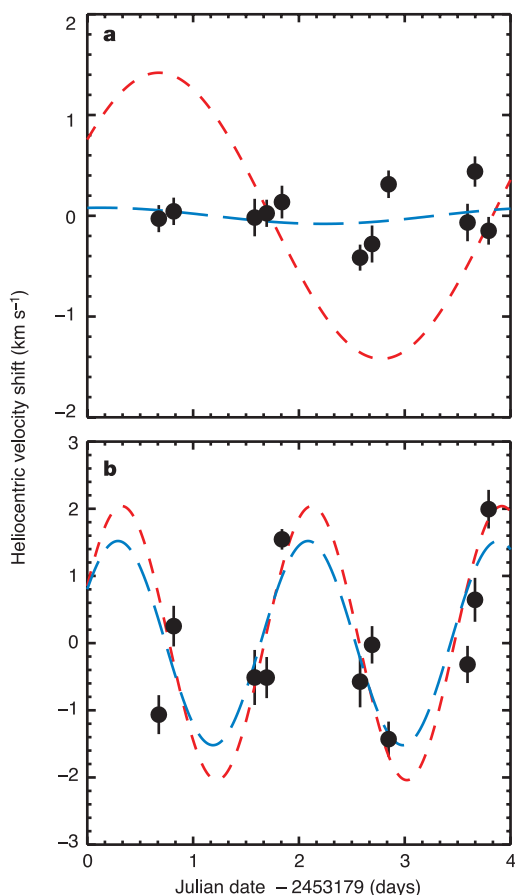


Figure 3 | RV measurements. RV variations (black dots) for SWEEPS-04 (a) and SWEEPS-11 (b) from VLT spectra. **a**, The best formal fit (blue curve) has insignificant amplitude, and at the 95% and 99.9% confidence levels we exclude companions more massive than $3.8 M_J$ and $5.3 M_J$. **b**, The best formal fit (blue curve) has an amplitude of $K = 1.5^{+0.9}_{-0.7} \text{ km s}^{-1}$ ($M_p = 9.7^{+5.6}_{-4.5} M_J$), formally yielding a 99% likelihood of detection, unless residual systematic errors are significant. Expected RV variations for a minimum-mass brown dwarf ($13 M_J$, dashed red curves) are also shown. The photometric and RV phases for SWEEPS-11 are consistent. Error bars indicate $\pm 1\sigma$.

below $V \sim 23$. The locations of the primaries found in our sample are thus consistent with the earlier findings that high-metallicity stars are more likely to host planets. (Although high metallicity provides a natural explanation for the redward shift at faint levels, we cannot exclude the alternative possibility that some of the faint hosts belong to the foreground disk. Proper motions of the host stars from future second-epoch observations will resolve this issue.)

The sample of RV-detected planets in the solar neighbourhood indicates that the frequency of occurrence of jovian planets is 5–10% for F–K dwarfs, about one-tenth of which are hot Jupiters with periods of less than 4.2 days; the frequency of planets around M dwarfs is $< 5\%$ (refs 4, 37). Our sample of candidates mainly belong to the Galactic bulge, the farthest such sample in the Galaxy, where the metallicity distribution is broader than in the solar neighbourhood^{21,22}. However, after taking into account the relation between planet frequency and metallicity in the local sample⁴¹ we would expect the overall planet frequency in the bulge to be similar to that in the solar neighbourhood. A major factor that reduces the completeness of transit surveys at longer periods is the lower geometric probability of a transit occurring, which is inversely proportional to the orbital radius¹⁹. After correcting for geometric transit probability and our detection efficiency (see Methods), we find that our 16 candidates (if all of them are assumed to be genuine planets) imply that about 0.42% of bulge stars more massive than $\sim 0.44 M_\odot$ are orbited by jovian planets with periods of less than 4.2 days. Because of the small-number statistics and uncertainties in the detection efficiencies, this fraction is uncertain by perhaps a factor of 2. Thus, within the statistical errors, the overall frequency of occurrence of planets derived from the SWEEPS data is consistent with that in the solar neighbourhood. For host stars more massive than $0.75 M_\odot$, the observed period distribution of the SWEEPS planets is also consistent with that in the solar-neighbourhood sample. However, for lower-mass stars the SWEEPS period distribution is systematically shifted to shorter periods.

Figure 4 plots the orbital periods against the host-star masses for the SWEEPS candidates, as well as for planets previously discovered

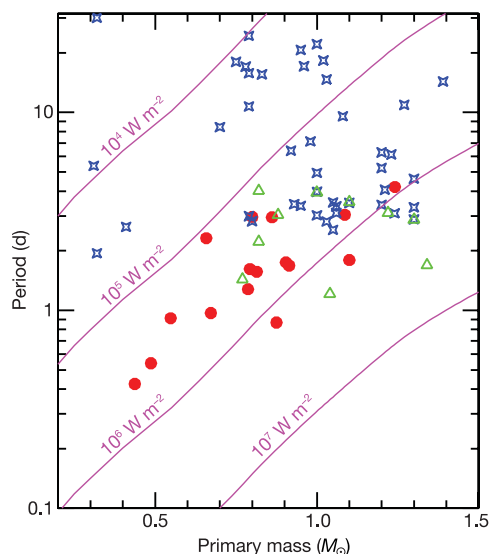


Figure 4 | Plot of orbital period against host-star mass. Orbital periods and host-star masses for extrasolar planets. Filled red circles are the 16 new SWEEPS candidates, green triangles are previously discovered transiting planets, and blue crosses are lower mass limits for planets detected through RV variability. The SWEEPS candidates extend the range of planetary orbital periods down to 0.42 days. Very few planets have irradiances above $2 \times 10^6 \text{ W m}^{-2}$ (which corresponds to an equilibrium temperature of 2,000 K) and none in the SWEEPS sample. The absence of USPPs around stars more massive than $0.9 M_\odot$ may be due to irradiative evaporation.

in RV and transit surveys. This figure shows that USPPs occur preferentially around low-mass stars. Thus, the reason that USPPs have not been revealed until now seems to be simply that previous transit and RV surveys have not reached down to sufficient numbers of low-mass, low-luminosity host stars with suitably high metallicity.

According to current models, Jupiter analogues with orbital periods of a few days cannot form *in situ* but must instead form at larger radii and then migrate inwards at an early stage by means of interaction with the circumstellar disk. Migration halts either when the disk dissipates or when the planet approaches the magnetically truncated inner edge of the disk at a few stellar radii^{43,44}. The existence of USPPs may imply that jovian planets around lower-mass stars can migrate to smaller orbital separations, perhaps because the size of the inner disk hole decreases with decreasing stellar radius.

To investigate the cause for the absence of USPPs around stars above $0.9 M_{\odot}$, we calculated the surface fluxes received at the planets using the known temperatures and radii of the host stars and the orbital separations. The diagonal lines in Fig. 4 mark loci of constant irradiation at the planetary surface. There are very few planets found at irradiances larger than $2 \times 10^6 \text{ W m}^{-2}$ (corresponding to an equilibrium surface temperature of $\sim 2,000 \text{ K}$ (ref. 45)), and none in the SWEPS sample. (It has been suggested⁴⁶ that the XUV and Lyman- α irradiation may be more important than the total irradiation in driving the escape of planetary atmospheres, but the former scales with the total irradiation times a factor related to the stellar Rossby number⁴⁷.) Fig. 4 thus indicates that planets that find themselves in such extreme thermal environments might be irradiatively evaporated^{46,48}, on a timescale that is less than $\sim 10 \text{ Gyr}$. In close proximity to less luminous stars, however, these ancient worlds have survived for more than twice the age of our own Solar System.

METHODS

Time-series and absolute photometry. During the 7-day SWEPS monitoring observations, we obtained 254 and 265 exposures in the V and I bands, respectively. Each HST orbit had a visibility period of 52 min, during which we typically obtained three exposures each in V and I. Exposure times in each filter were 339 s, giving a photometric accuracy per observation of about 0.003 mag at $V = 20$ and 0.04 mag at $V = 25$. A typical planetary transit lasts 1.5–3 h, so there are generally five to ten individual observations during each transit.

Our approach for time-series photometry is similar to that followed for the analysis of the 47 Tucanae data^{39,49}. The method relies on the formation of difference images, which requires the simultaneous determination of image-to-image spatial offsets and modelling of the signal response to spatial offsets and telescope focus changes. The resulting difference images contain noise due to poissonian photon statistics and charge-coupled device readout noise, plus genuine residuals due to stellar variability. Photometric time series are derived from the difference images by performing both aperture and point spread function (PSF)-fitting photometry, and on a star-by-star basis selecting the one providing the lower noise. The resulting time series are then cleaned of any remaining correlations with x, y position, to provide the final photometry (plus errors) for each target.

The CMD for the SWEPS field was determined from absolute photometry (Vegamag system) performed on twice-oversampled co-added images of the entire data set in V and I using the DAOPHOT II (ref. 50) PSF-fitting photometry package, with the photometric zero-points taken from the calibration work at the Space Telescope Science Institute.

Transit search and candidate selection. As described in the text, power spectra were calculated for each time series to detect the periodic transit signals. To evaluate the limiting S/N at which the number of false positive detections would be expected to be less than one, we performed a blind experiment with simulated transit signals. About 100 artificial planetary transits were blindly inserted into a simulated time-series data set of 115,000 non-variable stars with noise characteristics identical to those in the real SWEPS data. The data were then processed through our software to produce a list of transit candidates and S/N values as described above. All detected transits for which the calculated S/N exceeded 6.5 proved to be genuine events, with the first false positives appearing below this value. Hence a S/N limit of 6.5 was imposed to avoid any statistical false positives in our detected candidates.

Detection efficiency. Detection efficiency is defined as the fraction of actual transiting planets in the photometric data that are recovered by our search

technique. To estimate this efficiency, $\sim 80,000$ transits were introduced blindly into the actual time-series data, including appropriate noise. The simulated transit light curves assumed planet radii of $0.6\text{--}1.4 R_J$, a random period distribution over the interval $0.3\text{--}5.0$ days, and an appropriate range of transit durations. Our search technique was then used to produce a list of planetary candidates, including their physical characteristics. The results were compared with the input values to derive the detection efficiencies as a function of planetary radius, stellar magnitude and orbital period.

As expected, shorter periods are more easily detected than longer ones, and transits of brighter stars are more easily recovered than for fainter stars. At short periods of $0.5\text{--}1.5$ days, the detection efficiency is 90% at $V = 21$ for radii of $1.2 R_J$, 50% at $V = 22.7$, and 10% at $V = 26$. The corresponding stellar masses are $0.95 M_{\odot}$, $0.77 M_{\odot}$ and $0.46 M_{\odot}$, respectively. For $P = 2.5\text{--}3.5$ days and $R \sim 1.2 R_J$, the detection efficiency is 90% at $V = 20$, 50% at $V = 21.3$, and 10% at $V = 24$, corresponding to stellar masses of $1.08 M_{\odot}$, $0.88 M_{\odot}$ and $0.65 M_{\odot}$. Even at short periods and a planetary radius of $1.4 R_J$, the efficiency drops close to zero at $V = 26.5$, corresponding to a stellar mass of $0.42 M_{\odot}$.

There are $\sim 180,000$ stars brighter than $V = 27$ that form our targets for transit search. As stated above, the efficiency of planet detection drops to zero at $V = 26.5$, and our faintest candidate host has a brightness of $V = 26.2$.

Received 18 March 2005; accepted 10 August 2006.

- Mayor, M. & Queloz, D. A. Jupiter-mass companion to a solar-type star. *Nature* **378**, 355–359 (1995).
- Marcy, G. W. & Butler, R. P. A planetary companion to 70 Virginis. *Astrophys. J.* **464**, L147–L151 (1996).
- Schneider, J. The Extrasolar Planets Encyclopaedia (<http://www.obspm.fr/encycl/encycl.html>).
- Butler, R. P. et al. A Neptune-mass planet orbiting the nearby M dwarf GJ 4361. *Astrophys. J.* **617**, 580–588 (2004).
- Bonfils, X. et al. The HARPS search for southern extra-solar planets VI. A Neptune-mass planet around the nearby M dwarf Gl 581. *Astron. Astrophys.* **443**, L15–L18 (2005).
- Beaulieu, J.-P. et al. Discovery of a cool planet of 5.5 Earth masses through gravitational microlensing. *Nature* **439**, 437–440 (2006).
- Charbonneau, D., Brown, T. M., Latham, D. W. & Mayor, M. Detection of planetary transits across a Sun-like star. *Astrophys. J.* **529**, L45–L48 (2000).
- Henry, G., Marcy, G. W., Butler, R. P. & Vogt, S. S. A transiting 51 Peg-like planet. *Astrophys. J.* **529**, L41–L44 (2000).
- Udalski, A. et al. The Optical Gravitational Lensing Experiment. Additional planetary and low-luminosity object transits from the OGLE 2001 and 2002 observational campaigns. *Acta Astron.* **53**, 133–149 (2003).
- Konacki, M., Torres, G., Jha, S. & Sasselov, D. An extrasolar planet that transits the disk of its parent star. *Nature* **421**, 507–509 (2003).
- Bouchy, F. et al. Two new ‘very hot Jupiters’ among the OGLE transiting candidates. *Astron. Astrophys.* **421**, L13–L16 (2004).
- Moutou, C., Pont, F., Bouchy, F. & Mayor, M. Accurate radius and mass of the transiting exoplanet OGLE-TR-132b. *Astron. Astrophys.* **424**, L31–L34 (2004).
- Pont, F. et al. The ‘missing link’: A 4-day period transiting exoplanet around OGLE-TR-111. *Astron. Astrophys.* **426**, L15–L18 (2004).
- Bouchy, F. et al. Doppler follow-up of OGLE transiting companions in the Galactic bulge. *Astron. Astrophys.* **431**, 1105–1121 (2005).
- Alonso, R. et al. TrES-1: the transiting planet of a bright K0 V star. *Astrophys. J.* **613**, L153–L156 (2004).
- Sato, B. et al. The N2K Consortium. II. A transiting hot Saturn around HD 149026 with a large dense core. *Astrophys. J.* **633**, 465–473 (2005).
- Bouchy, F. et al. ELODIE metallicity-biased search for transiting hot Jupiters II. A very hot Jupiter transiting the bright K star HD189733. *Astron. Astrophys.* **444**, 115–119 (2005).
- McCullough, P. et al. A transiting planet of a Sun-like star. *Astrophys. J.* (in the press); preprint at (<http://www.arXiv.org/astro-ph/0605414>) (2006).
- Borucki, W. J. & Summers, A. L. The photometric method of detecting other planetary systems. *Icarus* **58**, 121–134 (1984).
- Zoccali, M. et al. Age and metallicity distribution of the galactic bulge from extensive optical and near-IR stellar photometry. *Astron. Astrophys.* **399**, 931–956 (2003).
- Fulbright, J. P., McWilliam, A. & Rich, R. M. Abundances of Baade’s Window giants from Keck/HIRES spectra: I. Stellar parameters and [Fe/H] values. *Astrophys. J.* **636**, 821–841 (2006).
- Kuijken, K. & Rich, R. M. Hubble Space Telescope WFPC2 proper motions in two bulge fields: kinematics and stellar population of the Galactic bulge. *Astrophys. J.* **124**, 2054–2066 (2002).
- Zoccali, M. et al. The initial mass function of the Galactic bulge down to $\sim 0.15 M_{\odot}$. *Astrophys. J.* **530**, 418–428 (2000).
- Brown, T. M. et al. Stellar cluster fiducial sequences with the Advanced Camera for Surveys. *Astron. J.* **130**, 1693–1706 (2005).
- VandenBerg, D. A., Bergbusch, P. A. & Dowler, P. D. The Victoria-Regina stellar models: evolutionary tracks and isochrones for a wide range in mass and

- metallicity that allow for empirically constrained amounts of convective core overshooting. *Astrophys. J. Suppl. Ser.* **162**, 375–387 (2006).
26. Castelli, F. & Kurucz, R. L. in *Modelling of Stellar Atmospheres* (eds Piskunov, N., Weiss, W. W. & Gray, D. F.) A20 (Proc. 210th Symp. IAU, Astronomical Society of the Pacific, San Francisco, 2003).
 27. Sumi, T. *et al.* The Optical Gravitational Lensing Experiment: catalogue of stellar proper motions in the OGLE-II Galactic bulge fields. *Mon. Not. R. Astron. Soc.* **348**, 1439–1450 (2004).
 28. Rich, R. M. & Origlia, L. The first detailed abundances for M giants in Baade's Window from infrared spectroscopy. *Astrophys. J.* **634**, 1293–1299 (2005).
 29. Kovacs, G., Zucker, S. & Mazeh, T. A box-fitting algorithm in the search for periodic transits. *Astron. Astrophys.* **391**, 369–377 (2002).
 30. Agol, E., Farmer, A. & Mandel, K. Analytic lightcurves for planetary transit searches. *Astrophys. J.* **580**, L171–L175 (2002).
 31. Baraffe, I., Chabrier, G., Barman, T. S., Allard, F. & Hauschildt, P. H. Evolutionary models for cool brown dwarfs and extrasolar giant planets. The case of HD 209458. *Astron. Astrophys.* **402**, 701–712 (2003).
 32. Burrows, A., Hubeny, I., Hubbard, W. B. & Sudarsky, D. Theoretical radii of transiting giant planets: the case of OGLE-TR-56b. *Astrophys. J.* **610**, L53–L56 (2004).
 33. Brown, T. M. Expected detection and false alarm rates for transiting Jovian planets. *Astrophys. J.* **593**, L125–L128 (2003).
 34. Duquennoy, A. & Mayor, M. Multiplicity among solar-type stars in the solar neighbourhood. II. Distribution of the orbital elements in an unbiased sample. *Astron. Astrophys.* **258**, 485–524 (1991).
 35. Pont, F. *et al.* A planet-sized transiting star around OGLE-TR-122. Accurate mass and radius near the hydrogen-burning limit. *Astron. Astrophys.* **433**, L21–L24 (2005).
 36. Marcy, G. Observed properties of exoplanets: masses, orbits, and metallicities. *Prog. Theor. Phys. Suppl.* **158**, 24–42 (2005).
 37. Pont, F. *et al.* Doppler follow-up of OGLE planetary transit candidates in Carina. *Astron. Astrophys.* **438**, 1123–1140 (2005).
 38. Pont, F. *et al.* Radius and mass of a transiting M dwarf near the hydrogen-burning limit. OGLE-TR-123. *Astron. Astrophys.* **447**, 1035–1039 (2006).
 39. Gilliland, R. L. *et al.* A lack of planets in 47 Tucanae from a Hubble Space Telescope search. *Astrophys. J.* **545**, L47–L51 (2000).
 40. Albrow, M. D. *et al.* The frequency of binary stars in the core of 47 Tucanae. *Astrophys. J.* **559**, 1060–1081 (2001).
 41. Fischer, D. A. & Valenti, J. The planet-metallicity correlation. *Astrophys. J.* **622**, 1102–1117 (2005).
 42. Valenti, J. A. & Fischer, D. A. Spectroscopic Properties of Cool Stars (SPOCS). I. 1040 F, G, and K dwarfs from Keck, Lick, and AAT Planet Search programs. *Astrophys. J. Suppl. Ser.* **159**, 141–166 (2005).
 43. Lin, D. N. C., Bodenheimer, P. & Richardson, D. C. Orbital migration of the planetary companion of 51 Pegasi to its present location. *Nature* **380**, 606–607 (1996).
 44. Trilling, D. *et al.* Orbital migration of giant planets: modeling extrasolar planets. *Astrophys. J.* **500**, 428–439 (1998).
 45. Grauer, A. D. & Bond, H. E. The precataclysmic nucleus of Abell 41. *Astrophys. J.* **271**, 259–263 (1983).
 46. Baraffe, I. *et al.* The effect of evaporation on the evolution of close-in giant planets. *Astron. Astrophys.* **419**, L13–L16 (2004).
 47. Vardavas, I. M. The dependence of the Rossby number and XUV-Ly α emission flux with age for solar-like G-type stars. *Mon. Not. R. Astron. Soc.* **363**, L51–L55 (2005).
 48. Lecavelier des Etangs, A., Vidal-Madjar, A., McConnell, J. C. & Hebrard, G. Atmospheric escape from hot Jupiters. *Astron. Astrophys.* **418**, L1–L4 (2004).
 49. Gilliland, R. L., Nugent, P. E. & Phillips, M. M. High-redshift supernovae in the Hubble Deep Field. *Astrophys. J.* **521**, 30–49 (1999).
 50. Stetson, P. B. in *Astronomical Data Analysis Software and Systems I* (eds Worrall, D. M., Biemesderfer, C. & Barnes, J.) 297–306 (ASP Conf. Series Vol. 25, Astronomical Society of the Pacific, San Francisco, 1992).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This article is based on observations made with the NASA/ESA Hubble Space Telescope, obtained at the Space Telescope Science Institute, which is operated by the AURA, Inc. under a NASA contract, and with the Very Large Telescope at the ESO, Paranal, Chile. We thank R. Gilliland for his generous contribution of time and efforts for this project, and D. Bradstreet, I. Jordan, G. Kovacs, D. VandenBerg and the late A. Lubenow for their help at various stages of the project.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to K.C.S. (ksahu@stsci.edu).

Genome-wide genetic analysis of polyploidy in yeast

Zuzana Storchová¹, Amanda Breneman^{1*}, Jessica Cande^{1*}, Joshua Dunn^{1*}, Kendra Burbank², Eileen O'Toole⁴ & David Pellman^{1,3}

Polyploidy, increased sets of chromosomes, occurs during development, cellular stress, disease and evolution. Despite its prevalence, little is known about the physiological alterations that accompany polyploidy. We previously described 'ploidy-specific lethality', where a gene deletion that is not lethal in haploid or diploid budding yeast causes lethality in triploids or tetraploids. Here we report a genome-wide screen to identify ploidy-specific lethal functions. Only 39 out of 3,740 mutations screened exhibited ploidy-specific lethality. Almost all of these mutations affect genomic stability by impairing homologous recombination, sister chromatid cohesion, or mitotic spindle function. We uncovered defects in wild-type tetraploids predicted by the screen, and identified mechanisms by which tetraploidization affects genomic stability. We show that tetraploids have a high incidence of syntelic/monopolar kinetochore attachments to the spindle pole. We suggest that this defect can be explained by mismatches in the ability to scale the size of the spindle pole body, spindle and kinetochores. Thus, geometric constraints may have profound effects on genome stability; the phenomenon described here may be relevant in a variety of biological contexts, including disease states such as cancer.

Polyploidy is frequent in nature. Polyploidy of whole organisms occurs in plants, fungi, invertebrates and lower vertebrates such as fish and amphibians. For reasons that are not clear, polyploidy is poorly tolerated in higher vertebrates^{1,2}. Polyploidization is also thought to occur frequently during evolution³. Genome sequencing suggests that many contemporary genomes, including genomes of higher vertebrates, evolved from ancient genome duplications: duplicated genomic segments that preserve gene order and orientation are readily apparent in many genomes from yeast to humans⁴. These findings are consistent with Ohno's hypothesis that gene duplications create substrates for evolutionary diversity⁵.

Polyploidy can also occur in a subset of cells within an otherwise euploid organism. In most contexts, cells become polyploid by a specialized mechanism known as the endocycle, which truncates the cell cycle and prevents cell division⁶. However, some cell types, such as the hepatocytes of the liver, are able to divide as polyploid cells⁷. In addition, many different cell types can become polyploid in pathological circumstances: in response to cellular stress, after viral infection, or in disease states such as cancer². Thus, polyploidy occurs only in certain cell types and physiological contexts and, at least during development, there are mechanisms that limit the division of polyploid cells.

There are several potential advantages to polyploidy. Polyploidy, with the attendant increase in cell size, might provide a metabolic benefit^{1,2}; for example, *Escherichia coli* grown in limiting nutrient conditions spontaneously become polyploid⁸. Gene redundancy might, in principle, protect polyploid cells from deleterious mutations¹, although, where tested, this does not translate into resistance to DNA-damaging agents^{9,10}. However, polyploidy can also have costs. In particular, newly formed polyploids frequently manifest defects in the maintenance of genomic stability. Tetraploid

budding yeast exhibits high rates of chromosome loss^{11,12} as well as increased levels of mutation and recombination events after treatment with DNA-damaging agents or microtubule-destabilizing drugs¹³. Chromosome mis-segregation due to abnormal mitoses is seen in a variety of tetraploid animal cell types that occur after cytokinesis failure or cell fusion^{2,14,15}. Genetically unstable tetraploid cells have been suggested to promote tumorigenesis^{2,15}, an idea that we have recently validated using a model of mammary epithelial tumorigenesis in the mouse¹⁴. The genomic instability of newly formed polyploid cells should accelerate the pace of generating new variants, raising interesting parallels between the role of polyploidy in the evolution of organisms and the role of polyploidy in the development of tumours².

Although genomic instability is commonly associated with polyploidy, the underlying mechanisms are not understood. More broadly, the extent of the physiological changes that accompany polyploidy is unknown. We previously identified a genetic phenomenon that provides an experimental avenue towards addressing these questions. We found that deletion of the gene encoding the microtubule/kinetochore-binding protein Bik1 results in 'ploidy-specific lethality': whereas *bik1Δ* haploids or diploids are viable, *bik1Δ* triploids have strongly reduced fitness, and *bik1Δ* tetraploids are inviable¹⁶. The phenomenon of ploidy-specific lethality demonstrates that marked physiological changes accompany polyploidy. As a first step towards defining the physiological changes that result from polyploidy, we designed a genome-wide strategy to identify ploidy-specific lethal functions comprehensively. The results of this screen reveal that tetraploid cells have a strong and specific requirement for a small group of genes involved in the maintenance of genome stability. The most dramatic defect in tetraploid cells is a defect in mitosis. We propose that this mitotic defect can be

¹Department of Pediatric Oncology, Dana-Farber Cancer Institute, ²Department of Physics, Harvard University, and ³Department of Pediatric Hematology/Oncology, Children's Hospital, Harvard Medical School, Boston, Massachusetts 02115, USA. ⁴The Boulder Laboratory for 3D Electron Microscopy of Cells, University of Colorado, Boulder, Colorado 80309, USA.

*These authors contributed equally to this work.

explained, at least in part, by mismatches in scaling the size of the spindle components, thus reducing the fidelity of chromosome segregation because of geometric constraints.

A genome-wide screen for ploidy-specific lethality

We developed a strategy to create tetraploid derivatives of the diploid homozygous gene deletion collection of budding yeast strains¹⁷. Gene replacement cassettes were designed to delete either the mating type *MATa* or *MATα* locus to convert non-mating *MATa/α* diploids into mating-competent *MATa/Δ* or *MATΔ/α* diploids from which we could then select tetraploids (Fig. 1a and Supplementary Note 1).

Using this strategy, from the collection of 4,766 viable diploid homozygous gene deletions¹⁷, we obtained 3,421 paired *MATa/Δ* or *MATΔ/α* deletion strains (71.8% of the viable yeast gene deletions) that were then analysed by pairwise mating (Fig. 1b). We identified 17 genes that are required for survival of yeast tetraploids (examples shown in Fig. 1c and Supplementary Table 1). The ploidy-specific lethality of tetraploids lacking these genes was verified independently by a plasmid shuffle strategy (see Supplementary Note 2). This demonstrates that the mutant strains score in our assay because of lethality or reduced fitness of tetraploids, not because of mating defects.

At least two classes of mutants could not be interrogated by our strategy: the essential genes and the mutants lacking genes for homologous recombination (due to inefficient replacement of the *MAT* locus¹⁸). We therefore used the plasmid shuffle strategy to construct tetraploid strains from these groups (Supplementary Note 2). For essential genes, we focused on genes that were functionally related to those identified from the genome-wide screen, thus our characterization of these genes is not comprehensive. We analysed temperature-sensitive mutants for growth defects at various intermediate temperatures (Fig. 1d). Combining these approaches, we identified 39 out of 3,540 genes that are required for survival of tetraploid yeast (Supplementary Tables 1 and 2 and Supplementary Data 1). Representative gene deletions were also tested for ploidy-specific lethal phenotype in different strain backgrounds; this analysis demonstrates that ploidy-specific lethality is a general phenomenon in *Saccharomyces cerevisiae* (Supplementary Fig. 1). For simplicity, we will refer to all mutations that either cause lethality or temperature-sensitive growth in tetraploids as ploidy-specific lethal mutations.

The strong requirement for certain genes in yeast tetraploids is unlikely to be due to ploidy-specific changes in gene expression. The comparison of triplicate samples of wild-type *MATa/α* diploids and

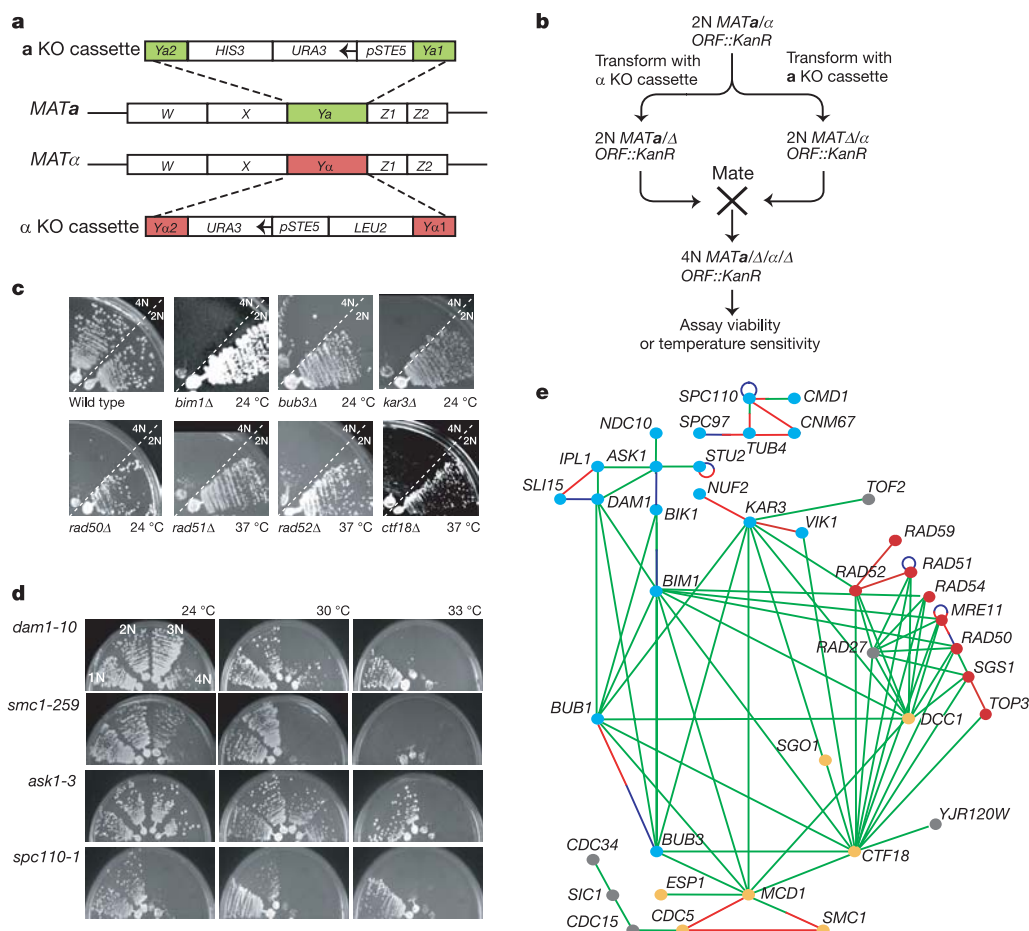


Figure 1 | A genome-wide strategy identifies a small subset of gene deletions/mutations that result in ploidy-specific lethality. **a**, Cartoon showing the structure of *MATa* and *MATα* replacement cassettes carrying selectable markers. Both cassettes also contain a reporter *URA3* gene under the control of the haplo-specific promoter *pSTE5*. Expression of *URA3* was thus turned on after successful replacement of the *MAT* locus, and then turned off after successful mating to produce tetraploids. KO, knockout. **b**, Flow diagram of the high-throughput screen. **c**, Ploidy-specific lethality of selected gene deletion strains identified in the high-throughput screen.

d, Examples of ploidy-dependent growth defects for strains with temperature-sensitive mutations in essential genes. **e**, Ploidy-specific lethality genes cluster into three functional groups: homologous recombination (red nodes), sister chromatid cohesion establishment and dissolution (yellow nodes), and mitotic spindle function (blue nodes); grey nodes mark genes that are not directly implicated in these functions. Green connectors, synthetic lethality; red connectors, protein-protein interactions; blue connectors, interactions by two-hybrid analysis. Created with the Osprey software.

MATa/a/α/α tetraploids did not reveal significant gene expression differences between diploids and tetraploids (Fig. 2, Supplementary Note 3 and Supplementary Data 2). This was confirmed by characterization of the steady-state protein levels for selected proteins (Supplementary Fig. 2). This result suggested that there must be other mechanisms underlying the observed specific changes in tetraploids and prompted our detailed comparison of genome stability in wild-type diploid and tetraploid cells.

Genomic stability in tetraploids

Notably, most of the ploidy-specific lethality genes fall into three functional groups: genes required for the repair of DNA damage by homologous recombination; genes required for sister chromatid cohesion establishment and dissolution; and a subgroup of genes required for the normal function of the mitotic spindle. Not only are almost all of these genes involved in the maintenance of genomic stability but there are extensive known genetic interactions among them¹⁹ (Fig. 1e). The functional links between these genes were further supported by our own genetic analysis (Supplementary Fig. 3). The remarkably small and specific group of ploidy-specific lethal mutations suggests that altered genomic stability is the major physiological change accompanying tetraploidization in budding yeast.

Next, we used genetic assays to systematically characterize genomic stability in wild-type tetraploids (Fig. 3a). By monitoring the segregation of a yeast artificial chromosome²⁰ we observed a 200-fold increase in the rate of chromosome loss in tetraploids relative to diploids. This is consistent with previous results demonstrating that chromosome mis-segregation is markedly frequent in yeast tetraploids^{11,12}. Notably, tetraploids exhibit an increase in the frequency of gene conversion between homologous chromosomes relative to diploids (assayed as recombination between two heteroalleles^{21,22}). Furthermore, a small but significant increase in the frequency of gross chromosomal rearrangements was observed in experiments that analysed rearrangements of markers within a yeast artificial chromosome²⁰. By contrast, the forward mutation rate (mutations within the *CAN1* gene²³) and rate of sister chromatid exchange (assayed as in ref. 25) are very similar in both diploid and tetraploid cells. Taken together, tetraploids exhibit massive

chromosome loss and an altered pattern of recombination relative to diploids, but do not display general 'hyper-recombination' and 'hyper-mutation' phenotypes¹⁸. The drug sensitivity profiles of tetraploids are consistent with these findings. Wild-type tetraploids have in comparison with diploids a five- to tenfold increased sensitivity to double-strand break-inducing agents and to the microtubule poison benomyl (Fig. 3b), but are not sensitive to osmotic stress, ultraviolet light, or oxidative damage (ref. 11 and data not shown).

Tetraploids require homologous recombination for survival

Haploid and diploid yeast, unlike animal cells, do not require homologous recombination for survival²⁴. The striking requirement

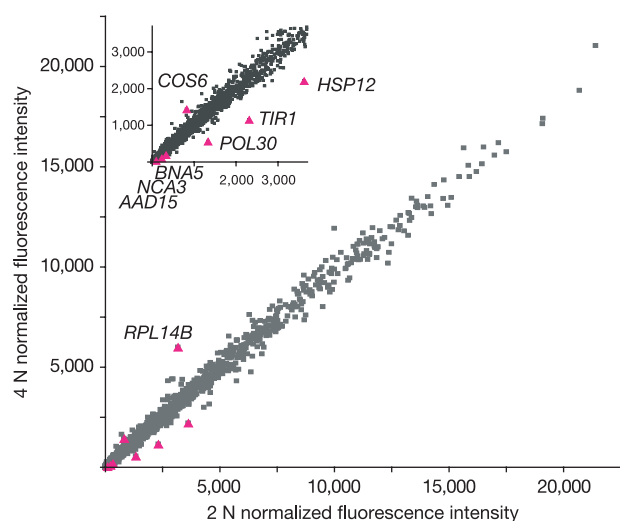


Figure 2 | The gene expression profile of diploid *MATa/a/α/α* and tetraploid *MATa/a/α/α* is not significantly altered by increased ploidy. Scatter plot of diploid and tetraploid data sets as a function of signal intensity. Eight genes showing statistically significant change are marked in the scatter plot. The four of them (*AAD15*, *POL30*, *NCA3* and *TIR1*) that show statistically significant changes above a twofold threshold may be false positives because they are in the low signal intensity (signal to noise) portion of the data set.

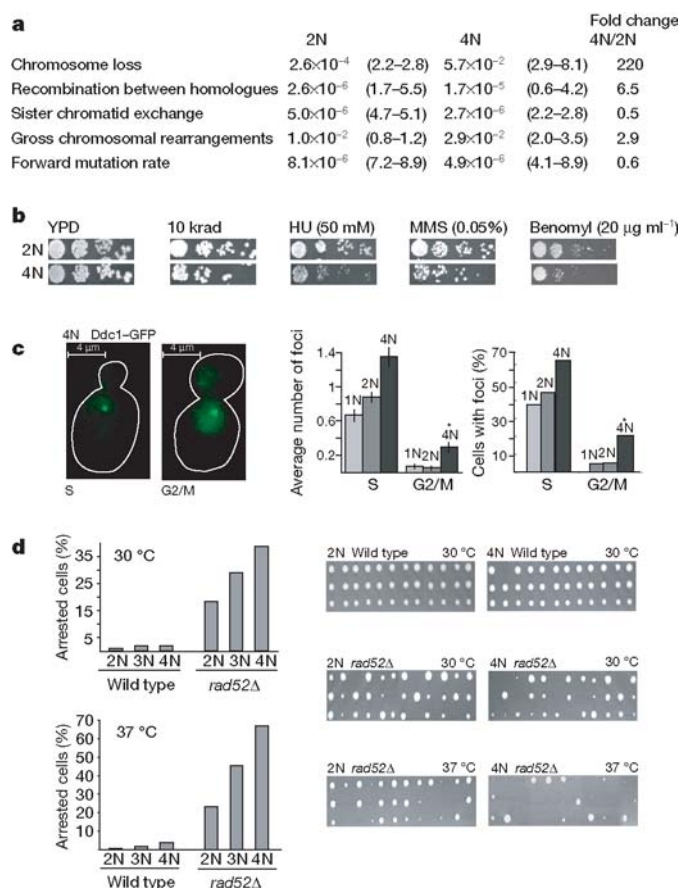


Figure 3 | Genome instability and requirement for homologous recombination in yeast tetraploids. **a**, Comparison of the rates and ranges of specific features of genomic stability in diploids and tetraploids (average of three experiments). The raw recombination rates between homologues almost certainly underestimate the true rates in wild-type tetraploids where all four loci are present. Tetraploids have twice as many chromosomes where a double-strand break can occur. There is a sixfold increase in the number of chromosome pairs capable of recombination. However, it is not clear whether the number of breaks or the number of donors will be rate-limiting for recombination during mitosis. **b**, Tetraploids show increased sensitivity to double-strand break inducing agents and to microtubule poisons. HU, hydroxyurea; MMS, methylmethane sulphonate. **c**, The number of DNA damage foci marked by Ddc1–GFP fusion protein foci increases with the ploidy. Tetraploids display an increase in Ddc1 foci in G2/M cells beyond that expected from the increase in ploidy. The differences are significant by χ^2 test ($P = 0.05$). Error bars indicate s.e.m. **d**, The fraction of small budded *rad52Δ* cells able to progress through the cell cycle diminishes in proportion with increasing ploidy. Small budded cells were micromanipulated and the number of arrested cells was scored after 12 h at a semi-permissive and non-permissive temperature. Images were taken after 3 days of incubation. A total of 104 individual cells were micromanipulated for each strain.

of tetraploids for homologous recombination genes could reflect elevated levels of spontaneous DNA lesions, especially given that the length of the cell cycle, and specifically S phase, is not increased with higher ploidy. One of the major sources of DNA lesions repaired by homologous recombination (mostly double-strand breaks) is DNA replication^{24,25}. One possibility is that an increased number of chromosome sets alone, without any defect in DNA replication, could create sufficient spontaneous DNA damage to trigger the strong requirement for homologous recombination.

To test this hypothesis, spontaneous DNA damage was visualized directly by imaging diploid and tetraploid cells expressing a Ddc1–GFP fusion protein²⁶. Ddc1 is a component of the 9-1-1 complex that is recruited early to DNA lesions. In haploids, Ddc1 foci appear during S phase and then disappear by the onset of anaphase^{26,27}. Consistent with our hypothesis, the number of S-phase cells (small budded cells) with Ddc1 foci increased in proportion to ploidy: a graded increase in 1N, 2N and 4N cells was observed for both the number of foci per cell and the percentage of cells with foci (Fig. 3c). This contrasted with mitotic cells where the number of Ddc1 foci in tetraploids was higher than expected based on the increase in chromosome content. This might indicate the persistence of DNA damage into mitosis (perhaps reflecting inefficient repair), additional DNA damage occurring in mitosis, or a subpopulation of cells transiently arrested to repair damage acquired during S phase.

To test further the idea that increased numbers of chromosome sets proportionally increase the amount of spontaneous DNA lesions, we characterized isogenic cells of varying ploidy that lacked Rad52, which is required for homologous recombination in budding yeast. A single double-strand break will activate the DNA damage checkpoint and will result in cell cycle arrest in *rad52Δ* cells¹⁸. The growth defect of *rad52Δ* strains might be explained if, during S phase, a fraction of cells acquire a spontaneous double-strand break that cannot be repaired. The tetraploid cells lacking Rad52 grow slowly at 24 °C and are inviable at 37 °C. Thus, increasing ploidy might proportionally increase the fraction of cells with spontaneous double-strand breaks, perhaps with most tetraploid cells at 37 °C acquiring a lesion that requires homologous recombination for repair. Consistent with this hypothesis, tetraploids lacking Rad52 arrested as large budded cells if shifted to 37 °C before S phase, but not if the temperature shift occurred after S-phase completion (Supplementary Fig. 4).

To test this hypothesis directly, small budded *rad52Δ* and control cells were separated by micromanipulation and monitored over time for progression through the cell cycle. At 30 °C, a temperature that is semi-permissive for *rad52Δ* tetraploids, 99% of wild-type 2N, 3N and 4N cells completed cell division within 4 h. By contrast, 17% of 2N, 28% of 3N and 38% of 4N *rad52Δ* cells underwent long-term cell cycle arrest, with morphology suggesting DNA damage checkpoint activation (Fig. 3d). Furthermore, at 37 °C, 24% of 2N, 44% of 3N and 69% of 4N *rad52Δ* cells underwent long-term cell cycle arrest (Fig. 3d). The linear increase in the fraction of arrested cells supports the idea that the number of cells with spontaneous DNA lesions increases proportionally with ploidy.

In agreement with this hypothesis, we did not detect defects in other processes that would aberrantly create double-strand breaks. Wild-type tetraploids maintained long telomeres over many generations, just like the diploid controls (Supplementary Fig. 5a). We did not observe any major defect in cell cycle progression and DNA replication in tetraploids by both flow cytometry and pulse-field gel electrophoresis of chromosomes isolated from synchronized cells (Supplementary Fig. 5b, c). We also did not observe an enhanced requirement for replication proteins such as Cdc7, Cdc6 or Pol32 in tetraploid cells (Supplementary Table 2 and Supplementary Data 1). Taken together, these results illustrate how a 'life or death' threshold can be crossed for a population of cells by increasing chromosome numbers. Thus, tetraploidy in yeast creates a situation, as observed in animal cells, where homologous recombination becomes essential for the completion of S phase^{24,28}.

Requirement of tetraploids for sister chromatid cohesion

Next, we characterized the requirement of tetraploid cells for sister chromatid cohesion, first determining whether wild-type tetraploid cells have a cohesion defect. Sister chromatid cohesion was monitored at two different locations on chromosome IV by a GFP fluorescence assay²⁹ in nocodazole-arrested cells. We observed that whereas wild-type diploid cells maintained sister chromatid cohesion during the arrest, tetraploid cells exhibited small but significant cohesion defects (about a threefold increase in the occurrence of sister chromatid separation during nocodazole arrest (Supplementary Fig. 6a). Notably, tetraploids arrested in mitosis as efficiently as diploids after exposure to nocodazole and maintained this arrest throughout the experiment (Supplementary Fig. 6b, c). Furthermore, haploid deletions of many of the identified ploidy-specific lethality genes (*KAR3*, *BIM1* and others) are also reported to have cohesion defects of a similar magnitude to those observed in wild-type tetraploid cells³⁰. The reason for the defect in sister chromatid cohesion in these haploid mutants is not known. We hypothesize, on the basis of the mitotic defects we describe below in wild-type tetraploids, that it is likely that similar mechanisms lead to cohesion defects in wild-type tetraploids and in these haploid mutants.

The mitotic defect of tetraploids

Our analysis of mitosis in wild-type tetraploids suggests that they have a specific defect in the mechanism of chromosome segregation. Relative to diploids, tetraploids have very high rates of chromosome loss, increased sensitivity to microtubule poisons, and an increased requirement for many proteins involved in spindle functions. However, cytoplasmic microtubule dynamics are not significantly different between diploids and tetraploids (Supplementary Fig. 7a) and no gene deletion affecting only cytoplasmic microtubule function (for example, *KAR9* or dynein) was identified as a ploidy-specific lethal mutation (Supplementary Data 1). Moreover, the dynamics of spindle microtubules, as measured by fluorescence recovery after photobleaching¹⁶, are not different between diploids and tetraploids (Supplementary Fig. 7b).

Direct observation of kinetochores in pre-anaphase triploid and tetraploid cells demonstrated that most kinetochores attach to microtubules (ref. 16 and Supplementary Fig. 7c). Consistent with this observation, tetraploids lacking the spindle checkpoint proteins Mad1, Mad2, or Mad3 are viable and there is no detectable mitotic delay for a population of wild-type tetraploids (Supplementary Fig. 5). By contrast, we observed a marked increase in the requirement for Ipl1, the yeast aurora B kinase, in tetraploid strains (Fig. 4a). Ipl1 promotes the dissolution of syntelic attachments in order to facilitate bipolar attachments³¹ and is also required for checkpoint activation triggered by kinetochores that are not under tension³². Moreover, tetraploids also require Bub1, Bub3 and Sgo1 for survival, all proteins implicated in responding to abnormal kinetochore–microtubule tension/attachment^{33,34}. These findings suggest that wild-type tetraploids could either have an increase in the number of syntelic attachments, triggering a high requirement for the Ipl1 pathway, or a primary defect in Ipl1 pathway function.

We monitored microtubule–kinetochore attachment using a non-replicating conditional di-centric minichromosome³⁵. After DNA replication, kinetochores of sister chromatids can either attach to microtubules from both poles (bi-orientation), where sister kinetochores are under tension, or they can both attach to a single pole (syntelic and monopolar attachment), where no tension is generated³⁴. Notably, like *IPL1*-deficient mutants, wild-type tetraploids display an increased frequency of syntelic/monopolar attachments in comparison to diploid controls (3.6-fold, Fig. 4b). Without correction, syntelic attachments lead to chromosome non-disjunction. We measured the frequency of non-disjunction by a visual assay following the distribution of a GFP-marked chromosome IV into daughter cells³⁶, and found that 2 out of 405 anaphase tetraploid cells had undergone non-disjunction, whereas no instances of non-disjunction

were observed in more than a thousand (0 out of 1,088) diploid anaphase cells. The frequency of non-disjunction detected by the visual assay closely matched the rates of chromosome loss (Fig. 3a). Thus, the high rate of chromosome loss in tetraploids may be largely explained by chromosome non-disjunction, which is a likely consequence of syntelic attachments.

Several lines of evidence support the idea that the observed increase in the frequency of kinetochores attached to only one pole is not due to decreased Ipl1 function. Ipl1 expression is not diminished in tetraploids either at the messenger RNA or protein level (Supplementary Fig. 2). In addition, we monitored the phosphorylation

of serine 10 on histone H3, an *in vivo* marker for Ipl1 kinase activity³⁷, in cells after release from a G1 block. The Ipl1-dependent phosphorylation of serine 10 on histone H3 is unchanged in tetraploid cells relative to the diploid control (Fig. 4c). This supports the conclusion that tetraploidy per se increases the frequency of syntelic/monopolar attachments.

Why is there an increase in syntelic/monopolar attachments in tetraploid cells? Because Ipl1 activity seems to be intact, we considered the possibility that scaling restrictions, with a resulting alteration in spindle geometry, might explain the chromosome segregation defect. Although cell volume increases in direct proportion to ploidy (Fig. 5a), some intracellular structures cannot be precisely scaled: individual kinetochores, although doubled in number (Supplementary Fig. 8), should be identical whether they are in a haploid or a tetraploid cell. Moreover, we observed that pre-anaphase spindle length is the same in haploids, diploids, triploids and tetraploids (Fig. 5a and ref. 16). By contrast, the disc-shaped spindle pole body (SPB) is assembled in a manner where doubling the number of components (expected in cells with double ploidy) roughly doubles its surface area and its nuclear microtubule nucleating capacity^{38,39}. Indeed, high-resolution

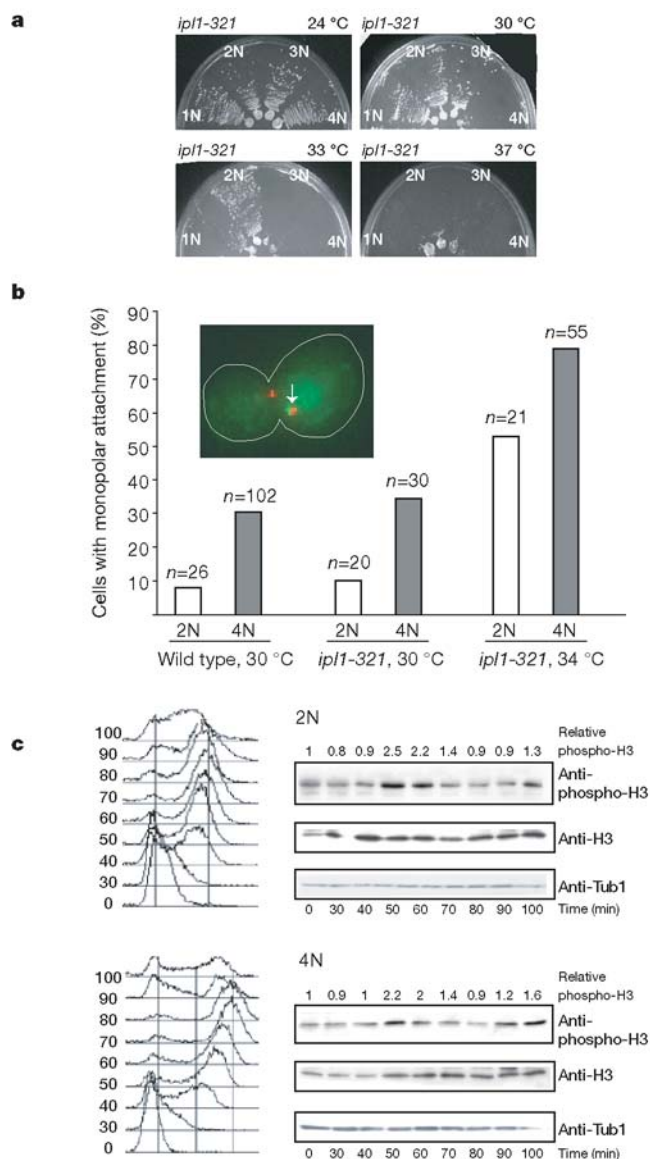


Figure 4 | Tetraploids have increased syntelic attachments without compromised Ipl1/aurora B activity. **a**, Tetraploidy enhances the growth defect in the *ip1-321* strain. **b**, Tetraploids show a 3.6-fold increase in the frequency of syntelic/monopolar attachments relative to diploids. Chromosome attachment was visualized by an assay using a conditional di-centric minichromosome⁴¹. The inset is an example of a cell with syntelic/monopolar attachment. Minichromosome is labelled green; SPBs are red. **c**, Phosphorylation of histone H3 by Ipl1 kinase during the cell cycle is not impaired in tetraploid cells—the normalized signal intensity of phosphorylated H3 approximately doubles during mitosis in both strains. Wild-type *MATa/a* and *MATa/a/a/a* strains were synchronized in G1 with α -factor and released into fresh media. Left panel shows DNA content by FACS.

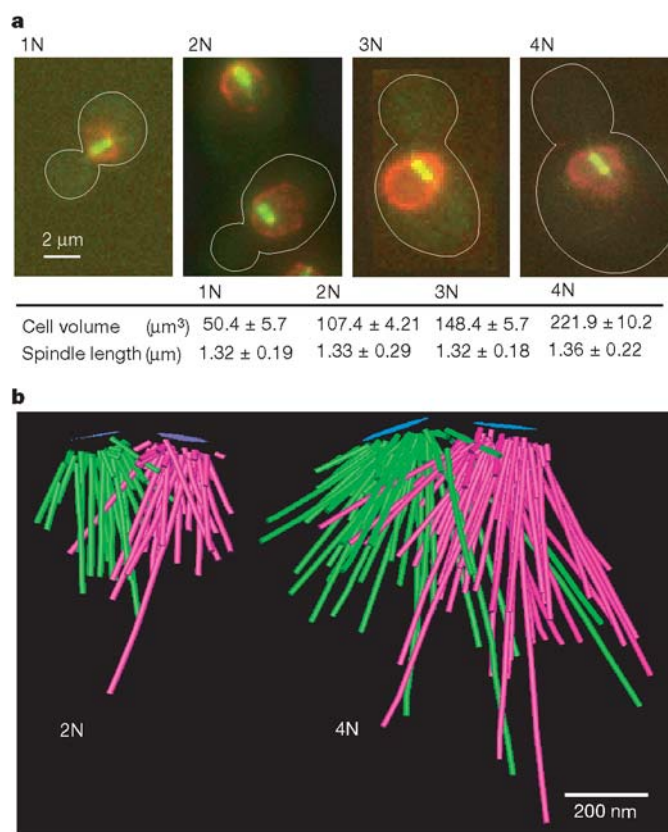


Figure 5 | Scaling effects can explain the increase in syntelic/monopolar attachments observed in tetraploids. **a**, Although cell volume increases proportionally with genome size, pre-anaphase spindle length does not. The nucleoporin Nup116 marking the nuclear envelope is in red (Nup116-YFP), Tub1-GFP, labelling the mitotic spindle, is in green. At least 100 cells was measured for each ploidy; error is indicated as s.e.m. **b**, Three-dimensional high-resolution electron tomography models of a forming spindle from a diploid and tetraploid cell. Microtubules nucleated from each spindle pole body are represented by the green and magenta lines, respectively. The positions of the SPB central plaques are indicated in purple and blue for 2N and 4N cells respectively. The four diploid SPBs measured contained 33, 37, 33 and 38 microtubules and the measured surface areas were 0.022, 0.020, 0.016 and 0.029 μm^2 . The four tetraploid SPBs measured contained 63, 66, 65 and 66 microtubules, and the measured surface areas were 0.044, 0.042, 0.045 and 0.042 μm^2 . See also Supplementary Movie 1. High-resolution electron tomography methods are described in Supplementary Notes 5.

electron tomography demonstrated that SPB nuclear surface area approximately doubles in tetraploids relative to diploids (mean $0.043 \pm 0.002 \mu\text{m}^2$ in tetraploids, $0.022 \pm 0.005 \mu\text{m}^2$ in diploids) and this increase is accompanied by a proportional increase in the number of nuclear microtubules (65 ± 1 microtubules per G1/S phase SPB in tetraploids and 35 ± 3 in diploids, Fig. 5b and Supplementary Movie 1). Thus, polyploidy in budding yeast creates mismatches between the size of the SPB, the length of spindle, and presumably the size and geometry of the kinetochore.

We suggest that this altered spindle geometry—the most striking difference we uncovered between diploid and tetraploid yeast—is likely to cause the increased frequency of syntelic attachment in tetraploids. For example, the expansion of SPB surface area in tetraploids may increase the fraction of microtubule pairs capable of forming syntelic attachments (Supplementary Fig. 9). This would be the case if microtubule pairs that originate close together from the nuclear surface of the SPB have a difficult time forming syntelic attachments, because they only connect to sister chromatids when there is bending between the kinetochores (Supplementary Fig. 9). Other geometry-based models that, for example, factor in the larger nuclear volume are also possible. Furthermore, sister chromatid cohesion is known to be required for normal kinetochore–microtubule attachment⁴⁰, and therefore the cohesion defect might contribute to the increased frequency of syntelic attachments as well as the high rate of chromosome loss observed in tetraploids. We note that the timing of cell cycle progression also does not scale with increased ploidy in budding yeast (Supplementary Fig. 5) or in animal cells (ref. 2). This may impose another type of scaling restriction that might affect the repair of spontaneous DNA damage (Fig. 3).

Discussion

Our study is the first comprehensive genetic analysis of polyploidy in any organism. We have identified a remarkably small and specific group of genes involved in the maintenance of genomic stability that is selectively required for viability/fitness of tetraploid yeast. The results suggest new hypotheses about how polyploidy affects cell physiology. We observed a markedly increased requirement for homologous recombination in tetraploids. This requirement does not seem to be due to a global defect in DNA replication and repair, but rather is more readily explained by a threshold effect due to the increase in spontaneous DNA damage associated with DNA replication of extra chromosome sets. Our genetic analysis together with the altered pattern of recombination in tetraploids suggests that tetraploids might have a defect in sister chromatid cohesion, which we validated experimentally. Finally, we uncovered a specific mitotic defect in tetraploid cells that might be explainable, at least in part, by geometric constraints in scaling the mitotic spindle. Together, these findings reveal mechanisms by which scaling genome and cell size can have profound effects on genomic stability. Such mechanisms might alter genome stability in various circumstances where newly formed polyploid cells are generated^{1,2}.

Our findings may have implications for human disease states associated with altered ploidy, most obviously cancer. We have recently demonstrated that tetraploidy can promote chromosomal instability and tumorigenesis in p53-null cells¹⁴. The mechanism by which tetraploidy results in chromosomal instability is not known, but the extra centrosomes that accompany tetraploidy after cell division failure have been assumed to have a major role. However, as in previous studies of tetraploid yeast^{12,16}, the tetraploid cells studied here were generated by mating. Thus, they have larger SPBs (centrosome equivalents)³⁸, but they do not contain extra SPBs/centrosomes. Therefore, increased ploidy per se, without extra centrosomes, can result in genomic instability.

The phenomenon of ploidy-specific lethality may also be useful for characterizing the physiology of tumour cells, if polyploid cells share common features with aneuploid, often near-triploid, tumour cells².

There are reasons to think that this could be the case. Mammalian CLIP-170, the orthologue of *BIK1*, a ploidy-specific lethality gene, is not essential in the mouse⁴¹, but RNA interference has revealed that it is essential in several cancer cell lines (ref. 42 and references therein). We found that tetraploid yeast have markedly increased requirements for genes involved in homologous recombination, and at least some tumours have been suggested to have an enhanced requirement for recombinational repair and elevated levels of recombination proteins^{43,44}. Finally, aurora B is overexpressed in many tumours, and several drugs targeting aurora B have been developed for use as chemotherapeutic agents⁴⁵. Our finding of an increased requirement for *Ipl1/aurora B* in tetraploid yeast potentially provides a rationale for why these drugs might be successful. Thus, the physiological changes identified by ploidy-specific lethality might be the basis for a strategy to identify new cancer therapeutics.

METHODS

Strain construction. All strains are derivatives of S288C (BY series, *leu2Δ*, *his3Δ*, *ura3Δ*, *lys2Δ/met15Δ*) or W303 (*ade2-101*, *ura3-52*, *trp1-1*, *his3-1,3*, *leu2-1,112*). The *MAT* locus replacement cassettes were constructed by three-step fusion PCR and subsequently verified by DNA sequencing. The cassettes were transformed using a modified LiAc transformation method. Further details are available in Supplementary Notes 1 and 2. Other strains used in this study are listed in Supplementary Table 3. The replacement cassettes, all strains, and information on their construction are available upon request.

Genetic analysis. Methylmethane sulphonate, hydroxyurea, γ -irradiation and benomyl sensitivity was determined with freshly made YPD plates containing the appropriate amount of the active agent. Overnight cultures of strains to be tested were diluted to equal cell numbers, and 5- μl aliquots of each dilution were plated. Growth was assessed after 1, 2, and 3 days at 30 °C. Recombination rates^{22,46}, chromosome loss rates, chromosomal rearrangement rates²⁰, and forward mutation rates²³ were determined by fluctuation tests using the median method⁴⁷ and were repeated three to five times for each genotype.

Imaging. Cells were observed in a fully automated Zeiss 200M inverted microscope (Carl Zeiss). Simultaneous photobleaching and fluorescence illumination was achieved using a beamsplitter (95% illumination/5% laser) to direct both the fluorescence light and the laser beam into the fluorescence light train of the microscope. All image acquisition, manipulations and fluorescence intensity measurements were performed using SlideBook. Measurements of FRAP were made as previously described¹⁶. The fluorescence intensity of a 2×2 pixel square ($0.2 \mu\text{m}^2$) was measured at each time point. The intensity measurements were corrected for both background fluorescence (using background values of the bleached cell) and loss of fluorescence that occurred during imaging (using average background fluorescence of 5–8 unbleached neighbour cells). The data were analysed using GraphPad Prism software. The high-resolution electron tomography is described in Supplementary Note 6.

Protein analysis. Lysates were isolated by 10 min incubation in 0.1 N NaOH followed by 10 min incubation in Laemli buffer at 96 °C. Standard SDS polyacrylamide gel electrophoresis and immunoblotting was performed. All antibodies used are commercially available with the exception of Clb2 and Ase1 antibody (see Supplementary Note 3).

Flow cytometry. Cultures were fixed in 70% ethanol, re-hydrated in 50 mM sodium citrate buffer and treated with RNase I at 37 °C overnight. After staining with propidium iodide ($16 \mu\text{g ml}^{-1}$ final concentration), cells were sonicated and flow cytometry was performed on a Becton Dickinson FACScan. For analysis we used Cell Quest software. For each experiment at least 10,000 cells were analysed.

Pulse-field gel electrophoresis. Agarose plugs containing cells from elutriated cultures were prepared according to standard protocol from the manufacturer (Bio-Rad). Electrophoresis was performed at 14 °C in a CHEF Mapper (Bio-Rad), using 1% agarose gels in 0.5× Tris-borate-EDTA, a voltage gradient of 6 V cm^{-1} , a switch angle of 120°, and switch times of 15 s (initial) to 60 s (final). Total run time was 24 h.

Received 24 May; accepted 17 August 2006.

1. Comai, L. The advantages and disadvantages of being polyploid. *Nature Rev. Genet.* **6**, 836–846 (2005).
2. Storchova, Z. & Pellman, D. From polyploidy to aneuploidy, genome instability and cancer. *Nature Rev. Mol. Cell Biol.* **5**, 45–54 (2004).
3. Otto, S. P. & Whitton, J. Polyploid incidence and evolution. *Annu. Rev. Genet.* **34**, 401–437 (2000).
4. Taylor, J. S. & Raes, J. Duplication and divergence: The evolution of new genes and old ideas. *Annu. Rev. Genet.* **38**, 615–643 (2004).

5. Ohno, S., Wolf, U. & Atkin, N. B. Evolution from fish to mammals by gene duplication. *Hereditas* **59**, 169–187 (1968).
6. Edgar, B. A. & Orr-Weaver, T. L. Endoreplication cell cycles: more for less. *Cell* **105**, 297–306 (2001).
7. Guidotti, J. E. *et al.* Liver cell polyploidization: a pivotal role for binuclear hepatocytes. *J. Biol. Chem.* **278**, 19095–19101 (2003).
8. Akerlund, T., Nordstrom, K. & Bernander, R. Analysis of cell size and DNA content in exponentially growing and stationary-phase batch cultures of *Escherichia coli*. *J. Bacteriol.* **177**, 6791–6797 (1995).
9. Mortimer, R. K. Radiobiological and genetic studies on a polyploid series (haploid to hexaploid) of *Saccharomyces cerevisiae*. *Radiat. Res.* **9**, 312–326 (1958).
10. Mable, B. K. & Otto, S. P. Masking and purging mutations following EMS treatment in haploid, diploid and tetraploid yeast (*Saccharomyces cerevisiae*). *Genet. Res.* **77**, 9–26 (2001).
11. Andalis, A. A. *et al.* Defects arising from whole-genome duplications in *Saccharomyces cerevisiae*. *Genetics* **167**, 1109–1121 (2004).
12. Mayer, V. W. & Aguilera, A. High levels of chromosome instability in polyploids of *Saccharomyces cerevisiae*. *Mutat. Res.* **231**, 177–186 (1990).
13. Mayer, V. W., Goin, C. J., Arras, C. A. & Taylor-Mayer, R. E. Comparison of chemically induced chromosome loss in a diploid, triploid, and tetraploid strain of *Saccharomyces cerevisiae*. *Mutat. Res.* **279**, 41–48 (1992).
14. Fujiwara, T. *et al.* Cytokinesis failure generating tetraploids promotes tumorigenesis in p53-null cells. *Nature* **437**, 1043–1047 (2005).
15. Nigg, E. A. Centrosome aberrations: cause or consequence of cancer progression? *Nature Rev. Cancer* **2**, 815–825 (2002).
16. Lin, H. *et al.* Polyploids require Bik1 for kinetochore-microtubule attachment. *J. Cell Biol.* **155**, 1173–1184 (2001).
17. Giaever, G. *et al.* Functional profiling of the *Saccharomyces cerevisiae* genome. *Nature* **418**, 387–391 (2002).
18. Paques, F. & Haber, J. E. Multiple pathways of recombination induced by double-strand breaks in *Saccharomyces cerevisiae*. *Microbiol. Mol. Biol. Rev.* **63**, 349–404 (1999).
19. Tong, A. H. *et al.* Systematic genetic analysis with ordered arrays of yeast deletion mutants. *Science* **294**, 2364–2368 (2001).
20. Huang, D. & Koshland, D. Chromosome integrity in *Saccharomyces cerevisiae*: the interplay of DNA replication initiation factors, elongation factors, and origins. *Genes Dev.* **17**, 1741–1754 (2003).
21. Kadyk, L. C. & Hartwell, L. H. Sister chromatids are preferred over homologs as substrates for recombinational repair in *Saccharomyces cerevisiae*. *Genetics* **132**, 387–402 (1992).
22. Klein, H. L. RDH54, a RAD54 homologue in *Saccharomyces cerevisiae*, is required for mitotic diploid-specific recombination and repair and for meiosis. *Genetics* **147**, 1533–1543 (1997).
23. Whelan, W. L., Gocke, E. & Manney, T. R. The *CAN1* locus of *Saccharomyces cerevisiae*: fine-structure analysis and forward mutation rates. *Genetics* **91**, 35–51 (1979).
24. Krogh, B. O. & Symington, L. S. Recombination proteins in yeast. *Annu. Rev. Genet.* **38**, 233–271 (2004).
25. Pan, X. *et al.* A DNA integrity network in the yeast *Saccharomyces cerevisiae*. *Cell* **124**, 1069–1081 (2006).
26. Melo, J. A., Cohen, J. & Toczyski, D. P. Two checkpoint complexes are independently recruited to sites of DNA damage *in vivo*. *Genes Dev.* **15**, 2809–2821 (2001).
27. Lisby, M., Barlow, J. H., Burgess, R. C. & Rothstein, R. Choreography of the DNA damage response: spatiotemporal relationships among checkpoint and repair proteins. *Cell* **118**, 699–713 (2004).
28. Sonoda, E. *et al.* Rad51-deficient vertebrate cells accumulate chromosomal breaks prior to cell death. *EMBO J.* **17**, 598–608 (1998).
29. Michaelis, C., Ciosk, R. & Nasmyth, K. Cohesins: chromosomal proteins that prevent premature separation of sister chromatids. *Cell* **91**, 35–45 (1997).
30. Mayer, M. L. *et al.* Identification of protein complexes required for efficient sister chromatid cohesion. *Mol. Biol. Cell* **15**, 1736–1745 (2004).
31. Tanaka, T. U. *et al.* Evidence that the Ipl1-Sli15 (Aurora kinase-INCENP) complex promotes chromosome bi-orientation by altering kinetochore-spindle pole connections. *Cell* **108**, 317–329 (2002).
32. Biggins, S. & Murray, A. W. The budding yeast protein kinase Ipl1/Aurora allows the absence of tension to activate the spindle checkpoint. *Genes Dev.* **15**, 3118–3129 (2001).
33. Indjeian, V. B., Stern, B. M. & Murray, A. W. The centromeric protein Sgo1 is required to sense lack of tension on mitotic chromosomes. *Science* **307**, 130–133 (2005).
34. Tanaka, T. U., Stark, M. J. & Tanaka, K. Kinetochore capture and bi-orientation on the mitotic spindle. *Nature Rev. Mol. Cell Biol.* **6**, 929–942 (2005).
35. Dewar, H., Tanaka, K., Nasmyth, K. & Tanaka, T. U. Tension between two kinetochores suffices for their bi-orientation on the mitotic spindle. *Nature* **428**, 93–97 (2004).
36. Cheeseman, I. M. *et al.* Implication of a novel multiprotein Dam1p complex in outer kinetochore function. *J. Cell Biol.* **155**, 1137–1145 (2001).
37. Hsu, J. Y. *et al.* Mitotic phosphorylation of histone H3 is governed by Ipl1/Aurora kinase and Glc7/PP1 phosphatase in budding yeast and nematodes. *Cell* **102**, 279–291 (2000).
38. Yoder, T. J., Pearson, C. G., Bloom, K. & Davis, T. N. The *Saccharomyces cerevisiae* spindle pole body is a dynamic structure. *Mol. Biol. Cell* **14**, 3494–3505 (2003).
39. Byers, B. & Goetsch, L. Duplication of spindle plaques and integration of the yeast cell cycle. *Cold Spring Harb. Symp. Quant. Biol.* **38**, 123–131 (1974).
40. Tanaka, T., Cosma, M. P., Wirth, K. & Nasmyth, K. Identification of cohesin association sites at centromeres and along chromosome arms. *Cell* **98**, 847–858 (1999).
41. Akhmanova, A. *et al.* The microtubule plus-end-tracking protein CLIP-170 associates with the spermatid manchette and is essential for spermatogenesis. *Genes Dev.* **19**, 2501–2515 (2005).
42. Green, R. A., Wollman, R. & Kaplan, K. B. APC and EB1 function together in mitosis to regulate spindle dynamics and chromosome alignment. *Mol. Biol. Cell* **16**, 4609–4622 (2005).
43. Raderschall, E. *et al.* Elevated levels of Rad51 recombination protein in tumor cells. *Cancer Res.* **62**, 219–225 (2002).
44. Flygare, J. *et al.* Effects of HsRad51 overexpression on cell proliferation, cell cycle progression, and apoptosis. *Exp. Cell Res.* **268**, 61–69 (2001).
45. Keen, N. & Taylor, S. Aurora-kinase inhibitors as anticancer agents. *Nature Rev. Cancer* **4**, 927–936 (2004).
46. Dong, Z. & Fasullo, M. Multiple recombination pathways for sister chromatid exchange in *Saccharomyces cerevisiae*: role of RAD1 and the RAD52 epistasis group genes. *Nucleic Acids Res.* **31**, 2576–2585 (2003).
47. Lea, D. E. & Coulson, C. A. The distribution of the numbers of mutants in bacterial populations. *J. Genet.* **49**, 264–285 (1949).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We are grateful to many colleagues from the yeast community for providing the reagents. We thank A. Amon, G. Fink, J. Haber, R. Rothstein, M. McLaughlin, M. Raschle and members of the Pellman laboratory for discussions; G. Fink, S. Elledge, J. Haber, A. Van Oudenaarden, M. McLaughlin, R. Rothstein and J. Walter for comments on the manuscript; C. Glavin for Supplementary Fig. 9; and M. Lenburg for guidance on the analysis of the expression profile data. D.P. was supported by an NIH grant and a gift from the G. Harold and Leila Y. Mathers Foundation. The Boulder Laboratory for 3D Electron Microscopy of Cells is supported by an NIH grant to J. R. McIntosh.

Author Information The transcriptional profiling data are available at MIAMExpress database (<http://www.ebi.ac.uk/miamexpress>) under accession number E-MEXP-822. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to D.P. (david_pellman@dfci.harvard.edu).

ARTICLES

AB₅ subtilase cytotoxin inactivates the endoplasmic reticulum chaperone BiP

Adrienne W. Paton¹, Travis Beddoe², Cheleste M. Thorpe³, James C. Whisstock², Matthew C. J. Wilce², Jamie Rossjohn², Ursula M. Talbot¹ & James C. Paton¹

AB₅ toxins are produced by pathogenic bacteria and consist of enzymatic A subunits that corrupt essential eukaryotic cell functions, and pentameric B subunits that mediate uptake into the target cell. AB₅ toxins include the Shiga, cholera and pertussis toxins and a recently discovered fourth family, subtilase cytotoxin, which is produced by certain Shiga toxigenic strains of *Escherichia coli*. Here we show that the extreme cytotoxicity of this toxin for eukaryotic cells is due to a specific single-site cleavage of the essential endoplasmic reticulum chaperone BiP/GRP78. The A subunit is a subtilase-like serine protease; structural studies revealed an unusually deep active-site cleft, which accounts for its exquisite substrate specificity. A single amino-acid substitution in the BiP target site prevented cleavage, and co-expression of this resistant protein protected transfected cells against the toxin. BiP is a master regulator of endoplasmic reticulum function, and its cleavage by subtilase cytotoxin represents a previously unknown trigger for cell death.

AB₅ toxins are key virulence factors of a range of bacteria, including Shiga toxigenic *Escherichia coli* (STEC), *Shigella dysenteriae*, *Vibrio cholerae*, enterotoxigenic *E. coli*, and *Bordetella pertussis*¹. These pathogens cause massive global morbidity and mortality, particularly among children in developing countries. STEC are a diverse group of *E. coli* strains defined by their ability to produce AB₅ toxins belonging to the Shiga toxin (Stx) family². They are an important cause of gastrointestinal disease in humans, particularly as these infections can result in life-threatening complications such as haemolytic uraemic syndrome (HUS). Much of the pathology associated with STEC disease is thought to be due to systemic effects of Stx. However, other factors are clearly important, and STEC strains differ markedly in their virulence in humans². Subtilase cytotoxin (SubAB) was discovered in a highly virulent O113:H21 STEC strain, which was responsible for an outbreak of HUS in South Australia in 1998, and has since been detected in several other STEC serotypes^{3,4}. SubAB is a potential bioterrorism agent: it is more toxic for Vero (African green monkey kidney) cells than Stx, and is lethal for mice, resulting in extensive microvascular damage, thrombosis and necrosis in several organs, including the brain, kidneys and liver³. It was named 'subtilase cytotoxin' because its 35-kDa A subunit (SubA) shares sequence homology with a subtilase-like serine protease of *Bacillus anthracis*³, although the sequence identity is low (26%). Subtilases are characterized by a conserved catalytic triad (Asp-His-Ser) and are found in a wide variety of microorganisms. However, none had previously been shown to have cytotoxic properties⁵. Mutagenesis of the crucial Ser₂₇₂ residue in SubA to Ala abolished *in vitro* and *in vivo* toxicity. Therefore, serine protease activity is central to the mechanism of action of the holotoxin³.

Identification of the SubAB substrate

To identify the crucial cellular substrate(s) for the toxin, we conducted proteomic analysis of Vero cell monolayers that had been treated for 3 h with 1 µg ml⁻¹ of either purified SubAB holotoxin or the inactive mutant SubA_{A272B} (see Supplementary Information).

Two-dimensional gel electrophoresis showed only four spots that were specific to the sample treated with the active toxin. These were excised, digested with trypsin and identified by matrix-assisted laser desorption/ionization–mass spectrometry (MALDI–MS) analysis as 14-3-3 gamma, 14-3-3 epsilon, mitochondrial matrix protein p32 and BiP (see Supplementary Table 1). Of these, only BiP (accession number P11021) showed a significant size discrepancy (~27 kDa on the gel compared with 72 kDa from the database) consistent with proteolytic cleavage by SubAB. Furthermore, the MALDI–MS analysis indicated that the tryptic peptides from the excised spot were derived exclusively from the carboxy-terminal domain of BiP (residues 448–654).

BiP (also known as GRP78) is a highly conserved endoplasmic reticulum (ER) chaperone that is essential for the survival of eukaryotic cells^{6–8}. It comprises an amino-terminal ATPase and a C-terminal protein-binding domain. One of its functions is to mediate the correct folding of nascent secretory proteins, by binding exposed hydrophobic domains at its C terminus; subsequent release is coupled to N-terminal domain-catalysed ATP-hydrolysis^{7,8}. BiP is also responsible for maintaining the permeability barrier of the ER membrane by sealing the luminal end of the Sec61 translocon pore, as well as for targeting terminally misfolded proteins to the Sec61 apparatus for degradation by the proteasome^{8,9}. BiP is vital for the unfolded protein response, acting as the ER stress-signalling master regulator⁷. It also opposes apoptosis by interfering with caspase activation. Therefore, disruption of BiP function has inevitably fatal consequences for the cell^{7,10–12}.

To confirm that SubAB mediates the cleavage of BiP, Vero cells were incubated with 1 µg ml⁻¹ purified SubAB or SubA_{A272B} for 3 h. Cell lysates were then subjected to SDS–polyacrylamide gel electrophoresis (SDS–PAGE) and western blot analysis using polyclonal goat anti-BiP antibodies (Fig. 1a). The BiP antibody labelled a 72-kDa species in the lysates of untreated Vero cells and those treated with the inactive toxin derivative. By contrast, the lysate of cells treated with active SubAB had a markedly diminished 72-kDa band as well as a

¹School of Molecular and Biomedical Science, University of Adelaide, South Australia 5005, Australia. ²Protein Crystallography Unit and ARC Centre of Excellence in Structural and Functional Microbial Genomics, Department of Biochemistry and Molecular Biology, Monash University, Clayton, Victoria 3800, Australia. ³Division of Geographic Medicine and Infectious Diseases, Tufts-New England Medical Center, Boston, Massachusetts 02111, USA.

new antibody-reactive species at ~28 kDa. This is consistent with the size of the C-terminal BiP fragment that was detected by proteomic analysis (~27 kDa; see Supplementary Table 1). The fact that only a single cleavage product was detected is consistent with the fact that the antibody (C-20) was raised against a 20-amino-acid peptide from the C terminus of BiP. An additional ~68-kDa species reacted weakly with the antibody, but was not susceptible to cleavage by SubAB. However, this species did not react with a separate polyclonal BiP antibody (N-20, Santa Cruz Biotech), and so is presumed to be a cross-reacting protein. BiP cleavage was evident within 20 min of treatment with $1 \mu\text{g ml}^{-1}$ SubAB and was maximal within 60 min (Fig. 1b). There was also clear evidence of BiP cleavage in Vero cells after 1 h at doses of SubAB as low as 0.32 ng ml^{-1} (Fig. 1c). Cleavage of the 72-kDa BiP species was completely blocked when Vero cells were pre-incubated for 30 min with $0.5 \mu\text{g ml}^{-1}$ brefeldin A, which inhibits retrograde transport by disrupting the Golgi apparatus (Fig. 1d). Co-localization of SubAB with BiP in toxin-treated Vero cells could also be demonstrated by confocal microscopy using Oregon Green-labelled SubAB and C-20 anti-BiP (Fig. 2). As BiP is located exclusively in the ER lumen⁸, co-localization with SubAB also confirms retrograde transport of the toxin to this compartment. Cleavage of BiP by SubAB was also demonstrated in human colonic epithelial (Hct8) and murine neuroblastoma (N2A) cell lines (result not shown, and Supplementary Fig. 2a, respectively).

Cleavage of BiP *in vivo*

To confirm that BiP is a target for SubAB *in vivo*, mouse liver tissue was examined 24 h after intraperitoneal injection with $5 \mu\text{g}$ SubAB. SDS-PAGE and western blot analysis with anti-BiP showed significant cleavage of BiP in toxin-treated mice, but not in control mice or in those injected with $5 \mu\text{g}$ SubA_{A272}B (Fig. 3a). Injection of SubAB, but not SubA_{A272}B, also resulted in an unfolded protein response (UPR) in the liver at 24 h, as judged by induction of the transcription factor CHOP (CCAAT/enhancer-binding protein

(C/EBP) homologous protein) (Fig. 3b). CHOP induction was also evident in N2A cells after 24 h treatment with $1 \mu\text{g ml}^{-1}$ SubAB or the known ER stressors tunicamycin or thapsigargin, but not after treatment with SubA_{A272}B or Shiga toxin (see Supplementary Fig. 2b).

In vitro cleavage of BiP

Finally, to confirm that SubAB is directly responsible for cleavage of BiP, purified bovine BiP was tested *in vitro* as a substrate for purified SubAB or SubA_{A272}B, as judged by SDS-PAGE (Fig. 3c). Untreated BiP migrated as a single 72-kDa species, and was not affected by exposure to SubA_{A272}B. However, in the presence of active SubAB, BiP was cleaved into 44-kDa and 28-kDa fragments. This confirms that SubAB directly cleaves BiP at a single position. Cleavage also occurred when BiP was treated with the purified catalytic subunit SubA, but it was not cleaved by purified SubB (Fig. 3c). The kinetics of BiP cleavage by SubAB and SubA were indistinguishable from each other (result not shown). The 28-kDa (presumptive C-terminal) BiP fragment was then excised from the gel and subjected to N-terminal sequence analysis. The first 10 residues had the sequence LDVCPLTLGI, which matches residues 417–426 of BiP. This indicates that cleavage occurs between two leucine residues at positions 416 and 417 in the 654-amino-acid BiP sequence; these are located in the exposed hinge that connects the ATPase and protein-binding domains¹³.

Most subtilisin-like enzymes are non-specific proteases that show

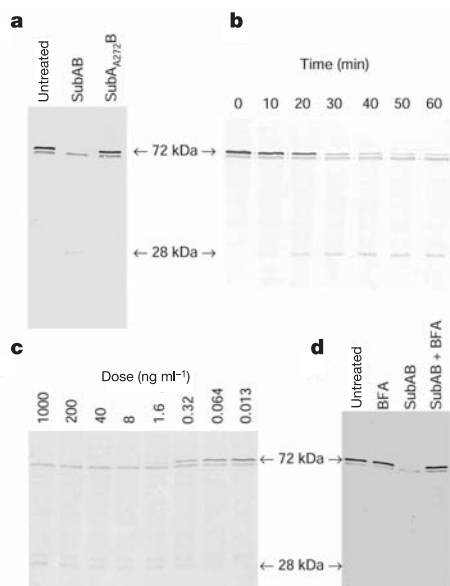


Figure 1 | SubAB cleaves BiP in Vero cells. Vero cell monolayers in 24-well plates were incubated at 37 °C in DMEM growth medium with or without purified SubAB or SubA_{A272}B, as indicated. Cell monolayers were then solubilized and subjected to SDS-PAGE and western blot analysis using polyclonal goat anti-BiP (see Methods). **a**, Cleavage of BiP by $1 \mu\text{g ml}^{-1}$ SubAB or SubA_{A272}B after 3 h incubation. **b**, Time course of BiP cleavage at $1 \mu\text{g ml}^{-1}$ SubAB. **c**, Dose-response experiment; cells were treated with the indicated dose of SubAB for 60 min. **d**, Effect of brefeldin A (BFA) on BiP cleavage; cells were pre-treated for 30 min with or without $0.5 \mu\text{g ml}^{-1}$ BFA, and then treated with or without $1 \mu\text{g ml}^{-1}$ SubAB for 60 min.

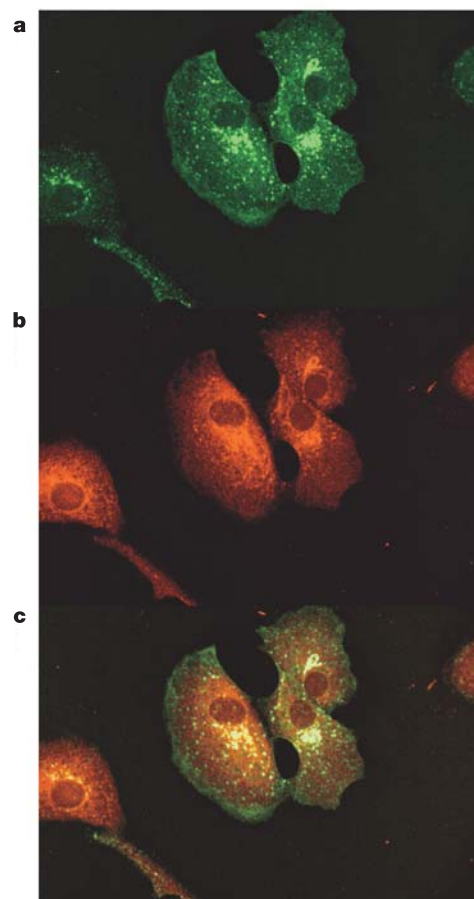


Figure 2 | Co-localization of SubAB and BiP. Vero cells growing on coverslips were treated with $1 \mu\text{g ml}^{-1}$ Oregon Green-SubAB for 60 min. Excess toxin was then removed and cells were incubated in fresh medium for 30 min. Cells were then fixed, permeabilized and reacted with goat anti-BiP (C-20), followed by Texas Red- (TR)- conjugated rabbit anti-goat IgG. Washed coverslips were examined by confocal microscopy. **a**, Oregon Green channel. **b**, Texas Red channel. **c**, Overlaid images.

a preference to cleave after hydrophobic residues. However, our proteomic data indicated that SubAB targeted a single protein in mammalian cells. Moreover, SubAB did not cleave purified human Hsp70 and Hsc70 proteins *in vitro* (see Supplementary Fig. 3a). These two members of the Hsp70 family are the closest homologues of BiP, and are identical at 7 of 11 amino-acid positions flanking the cleavage site, including the di-leucine motif.

Structural analysis of SubA

To understand the unique substrate specificity of SubA, we determined its crystal structure to 1.8 Å resolution (see Supplementary Information). The structure confirmed the homology to other members of the subtilisin S8A superfamily. A superposition with subtilisin Carlsberg (root mean squared deviation (r.m.s.d.) 1.63 Å over 239 C α atoms) revealed several features that account for its restricted substrate specificity. In particular, the catalytic His-Asp-Ser triad lies at the bottom of an uncharacteristically deep cleft formed by an insertion (81–88) at the N terminus of a helix that flanks the active site and the 234–239 loop on the opposite side of the cleft (Fig. 4). Analysis of the active-site cleft of SubA revealed a hydrophobic S1 pocket, consistent with this enzyme's cleavage of a Leu-Leu motif in BiP. Comparison with other subtilisin-like proteases, however, reveals that the S' side of the active site is restricted; movement of the 152–156 loop and rearrangement of the 119–123 loop creates a deep channel that leads out of the active site (Fig. 4). This mechanism of engendering substrate specificity in SubA, by the steric effect of loops occluding an active site, has been observed previously in thrombin, a chymotrypsin-like serine protease with exquisite substrate specificity¹⁴.

Mutagenesis of the target site in BiP

Identification of the cleavage site provided an opportunity to prove that the mechanism of cytotoxicity of SubAB is BiP cleavage. If a single amino-acid substitution at this site rendered the modified BiP resistant to cleavage by the toxin, co-expression of such a protein in transfected Vero cells might protect them from SubAB toxicity. A

full-length murine cDNA encoding for BiP was cloned into a prokaryotic expression vector (pBC) and mutated to encode a BiP derivative with a L₄₁₆ → D substitution (see Methods). French pressure cell lysates of recombinant *E. coli* expressing either native or mutated BiP (designated BiP_{D416}) were treated with 1 $\mu\text{g ml}^{-1}$ SubAB or SubA_{A272}B, and cleavage of BiP was assessed by western blot analysis using C-20 anti-BiP. The native BiP clone lysate contained a 72-kDa immunoreactive species, which was susceptible to cleavage by SubAB, but not SubA_{A272}B. In contrast, the BiP_{D416} clone lysate contained a 72-kDa immunoreactive species that was unaffected by either active or mutant toxin (Fig. 5a). Resistance of BiP_{D416} to cleavage was confirmed by exposing purified protein to purified SubAB *in vitro* (see Supplementary Fig. 3b).

Next, Vero cells were transfected with BiP or BiP_{D416} coding sequences inserted into the eukaryotic expression vector pEFIREP (see Methods). Western blot analysis indicated that BiP was almost completely digested when either Vero cells or those transfected with pEFIREP-BiP were exposed to 1 $\mu\text{g ml}^{-1}$ SubAB for 2 h. However, in cells transfected with pEFIREP-BiP_{D416}, some of the BiP remained intact after SubAB treatment, consistent with low-level co-expression of the resistant BiP_{D416} protein (Fig. 5b). These cells showed negligible cytotoxicity when exposed to 0.1 $\mu\text{g ml}^{-1}$ SubAB for 3 days, whereas untransfected Vero cells or those transfected with pEFIREP-BiP showed profound toxicity when exposed to the toxin (Fig. 5c). We also performed quantitative titrations of purified

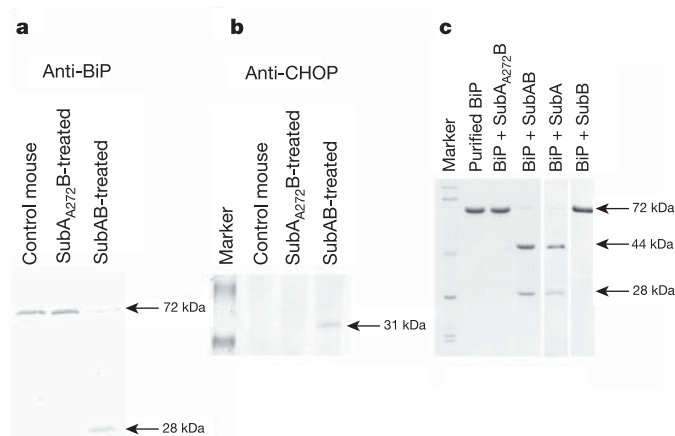


Figure 3 | *In vivo* cleavage of BiP and induction of CHOP, and *in vitro* cleavage of purified BiP. Liver tissue from a normal mouse, or from mice 24 h after intraperitoneal injection with 5 μg SubAB or SubA_{A272}B, was separated by SDS-PAGE and subjected to western blot analysis using: **a**, polyclonal goat anti-BiP (see Methods), or **b**, monoclonal anti-CHOP (9C8; Alexis Corporation, Lausen, Switzerland). The expected mobility of CHOP (31 kDa) is indicated; marker track shows 38-kDa and 26-kDa bands of Benchmark Prestained Protein Ladder. **c**, *In vitro* cleavage of purified BiP. Purified bovine BiP (5 μg) (Sigma) was incubated in DMEM with or without 100 ng SubAB, SubA_{A272}B, SubA alone or SubB alone, for 1 h at 37 °C in a volume of 15 μl . Proteins were then separated by SDS-PAGE and stained with Coomassie blue. Fragment sizes were determined with reference to the markers (BioRad 'broad range').

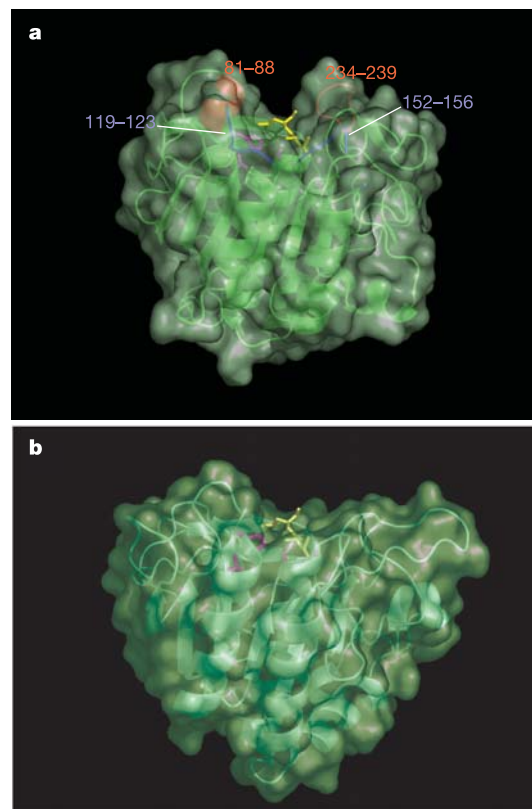


Figure 4 | Structural comparison of SubA with subtilisin Carlsberg.

a, Cartoon representation of SubA viewing from the S' end of the active site. The catalytic triad is in magenta stick and the position of a modelled P1 Leu/P1' Leu substrate is shown in yellow stick. The 234–239 loop and the N terminus of the helix flanking the active site (81–88) are highlighted in red. The elements of secondary structure lining the channel leading out of the S' side of the active site are highlighted in dark blue (119–123 and 152–156). **b**, Analogous view of the active site of subtilisin Carlsberg (PDB ID1C3L), with active-site residues in magenta. The position of a modelled P1/P1' substrate is shown in yellow. Note the relatively open active site in comparison with SubA.

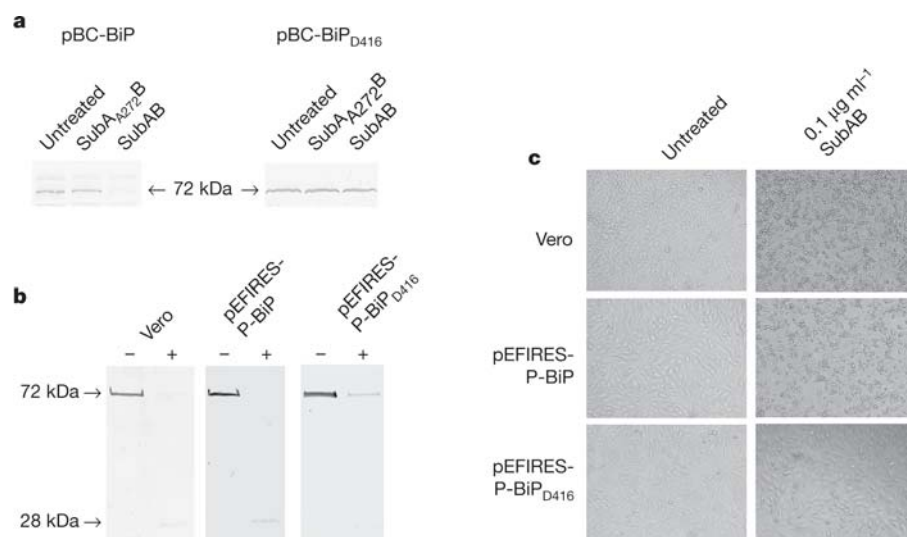


Figure 5 | Expression of SubAB-resistant BiP_{D416} prevents cytotoxicity. **a**, SubAB cleavage of BiP and BiP_{D416} expressed by recombinant *E. coli*. Lysates of *E. coli* JM109 carrying pBC-BiP or pBC-BiP_{D416} were treated with or without 1 μg ml⁻¹ SubAB or SubA_{A272}B for 1 h at 37 °C, and then subjected to SDS-PAGE and western blot analysis using polyclonal goat anti-BiP. The mobility of undigested BiP (72 kDa) is indicated. **b**, Cleavage of

BiP in transfected Vero cells. Untransfected Vero cells, or cells transfected with pEFIREs-P-BiP or pEFIREs-P-BiP_{D416}, were treated with (+) or without (-) 1 μg ml⁻¹ SubAB for 2 h at 37 °C, and then subjected to SDS-PAGE and western blot analysis as above. **c**, Cytotoxicity of SubAB for transfected cells. Cell monolayers were exposed to 0.1 μg ml⁻¹ SubAB and cytotoxicity was assessed after 3 days at 37 °C.

SubAB on Vero cells transfected with or without pEFIREs-P-BiP or pEFIREs-P-BiP_{D416}. There was no difference in endpoint titre when SubAB was assayed on untransfected Vero cells or those transfected with the native BiP construct (200,000 CD₅₀ per μg toxin). However, when assayed on Vero cells transfected with the BiP_{D416} construct, SubAB cytotoxicity was below the detectable threshold even at the maximum toxin concentration tested, and the 'endpoint titre' was by definition <50 CD₅₀ per μg toxin. Confocal microscopy using Oregon Green-labelled SubAB indicated that there was no defect in uptake or intracellular trafficking of the toxin in the transfected cells (result not shown). Therefore, resistance to SubAB cytotoxicity is a direct consequence of co-expression of BiP_{D416}.

Discussion

The findings of this study indicate that SubAB is unique in two ways. Unlike other bacterial AB₅ cytotoxins, it does not need to be translocated from the ER lumen into the cytosol in order to interact with its cellular target. For Ctx and Stx, this retro-translocation is achieved through subversion of the Sec61 translocon apparatus, which is normally used to transport terminally misfolded host proteins from the ER lumen into the cytosol for degradation, a process in which BiP is also involved^{15,16}. Most importantly, however, no other cytotoxin has been shown to target chaperone proteins or components of the ER directly. Defects in chaperone function, particularly those that affect ER stress responses, have been implicated in cellular senescence and a range of degenerative conditions including cataracts and Parkinson's and Alzheimer's diseases^{6,17}. Through its ability to rapidly and specifically abolish BiP function, SubAB provides a new tool in cell biology that enables *in vitro* modelling of key events in the pathogenesis of these diseases. BiP is a highly conserved master regulator of ER function, and is essential for survival of eukaryotes from simple yeasts to higher organisms such as mammals^{6,7}. There is negligible sequence variation in the vicinity of the cleavage site, indicating that SubAB has the potential for toxicity against a broad spectrum of life forms. Our findings reveal a previously undescribed mechanism of inducing cell death, through the specific targeting of a protease to disable chaperone function.

METHODS

Purified proteins. SubAB, its non-toxic derivative SubA_{A272}B and SubA were

purified by Ni-NTA chromatography as described previously^{3,18}. Purified bovine BiP was obtained from Sigma.

Cell culture and cytotoxicity assays. Vero cells were routinely grown in DMEM medium. For cytotoxicity assays, we washed confluent monolayers in 96-well flat-bottom trays with PBS, treated them with 50 μl of toxin (serially diluted in DMEM without FCS), and incubated them at 37 °C for 30 min. We then added 150 μl of medium supplemented with 2% FCS per well. Cytotoxicity was assessed microscopically after 3 days of incubation at 37 °C. We defined the toxin titre as the reciprocal of the maximum dilution that produced a cytopathic effect (rounding and detachment) on at least 50% of the cells in each well (CD₅₀ ml⁻¹)³.

Proteomic analysis. Proteomic analysis was carried out as described in Supplementary Information. N-terminal amino-acid sequence analysis was performed at the Australian Proteome Analysis Facility, Macquarie University, Sydney, Australia.

Western blot analysis. We separated proteins by SDS-PAGE and transferred them onto nitrocellulose as described previously^{19,20}. Filters were probed with polyclonal goat anti-BiP (C-20, Santa Cruz Biotech) (diluted 1:5000), followed by rabbit anti-goat IgG-alkaline phosphatase conjugate (BioRad Laboratories). Labelled bands were visualized using NBT/X-phosphate substrate (Roche Molecular Diagnostics). We used the Benchmark Prestained Protein Ladder (Invitrogen) as a size marker.

Confocal microscopy. Purified SubAB was labelled with Oregon Green (OG) using a Molecular Probes FluoReporter Oregon Green 488 Protein Labelling Kit (Invitrogen). This labelling had no discernible effect upon specific cytotoxicity for Vero cells. Vero cells grown on 1-cm coverslips in 24-well plates were treated with 1 μg ml⁻¹ OG-SubAB in DMEM for 60 min, then washed with PBS and incubated for a further 30 min in fresh DMEM. We then washed the cells with PBS, fixed them with 4% formaldehyde in PBS for 10 min and permeabilized them with 0.1% Triton X-100. Coverslips were then washed in PBS, blocked with 20% FCS in PBS for 1 h at 37 °C, and then treated with goat anti-BiP (C-20) (diluted in PBS/10% FCS) for 2 h at 37 °C. After three washes with PBS, the coverslips were reacted with Texas Red- (TR)- conjugated rabbit anti-goat IgG (Invitrogen) diluted in PBS/10% FCS for 1 h at 37 °C. Cells on the coverslips were then washed three times with PBS, twice with water, dried and mounted on glass slides using 3 μl of ProLong Gold Antifade (Invitrogen), allowed to cure for 1–2 h, and then sealed with nail varnish. We examined the cells using a BioRad MRC 600 Dual-Fluorescence Confocal microscope, captured images using CoMOS software (BioRad) and performed image overlays using Adobe Photoshop.

DNA methods. Routine DNA manipulations were carried out essentially as described previously²¹. DNA sequencing used dye-terminator chemistry and an ABI 3700 sequencer.

Site-directed mutagenesis of BiP. A 2.0-kb fragment containing the murine BiP coding sequence was excised from a full-length cDNA clone in the eukaryotic

expression vector pCMV4 (ref. 22) (a gift from M.J. Gething, University of Melbourne) using *Xba*I, inserted into the prokaryotic expression vector pBC (Stratagene) downstream of the vector *lac* promoter and ribosome binding site, and transformed into *E. coli* JM109 (ref. 23). We confirmed expression of BiP by one of the resultant recombinant *E. coli* transformants using western blot analysis. We then used overlap extension PCR mutagenesis to construct a derivative encoding a BiP protein with a L₄₁₆ → D substitution next to the SubAB cleavage site. This involved two separate PCR amplifications of pBC-BiP DNA template, using the Expand High Fidelity PCR Kit (Roche Molecular Diagnostics). The first used the universal M13 Forward primer and primer L416DOLR (5'-ACAAACATCAAGGTCTACCAGATCACCTGTATCC-3'). The second PCR used primer L416DOLF (5'-GGTGATCTGGTAGACCTTGATGTTTGTCCCTTAC-3') and the universal M13 Reverse primer. The two separate PCR products were then purified and combined, and the complete BiP coding region was reamplified using the M13 Forward and Reverse primers. The resultant PCR product was then cloned into pBC downstream of the vector *lac* promoter, and transformed into *E. coli* JM109. We purified recombinant plasmids from the resultant transformants and subjected them to sequence analysis to confirm that the correct mutation had been introduced. We also confirmed expression of a BiP protein using western blot analysis.

Expression of BiP in transfected Vero cells. The native BiP and BiP_{D416} coding sequences were excised from pBC using *Xba*I, inserted into the eukaryotic expression vector pEFIREP (ref. 24), and transformed into *E. coli* JM109. Purified plasmid DNA was extracted from transformants and subjected to restriction analysis to confirm correct orientation with respect to the vector promoter. We then transfected Vero cells with plasmid DNA using Lipofectamine 2000 reagent as recommended by the supplier (Invitrogen). Transfected cells were selected using 5 µg ml⁻¹ puromycin (Sigma).

Structural analysis. Details of conditions for crystallization of SubA and structure determination are provided in Supplementary Information.

Received 23 May; accepted 1 August 2006.

- Fan, E., Merritt, E. A., Verlinde, C. L. M. J. & Hol, W. G. J. AB₅ toxins: structures and inhibitor design. *Curr. Opin. Struct. Biol.* **10**, 680–686 (2000).
- Paton, J. C. & Paton, A. W. Pathogenesis and diagnosis of Shiga toxin-producing *Escherichia coli* infections. *Clin. Microbiol. Rev.* **11**, 450–479 (1998).
- Paton, A. W., Srimanote, P., Talbot, U. M., Wang, H. & Paton, J. C. A new family of potent AB₅ cytotoxins produced by Shiga toxinogenic *Escherichia coli*. *J. Exp. Med.* **200**, 35–46 (2004).
- Paton, A. W. & Paton, J. C. Multiplex PCR for direct detection of Shiga toxinogenic *Escherichia coli* producing the novel subtilase cytotoxin. *J. Clin. Microbiol.* **43**, 2944–2947 (2005).
- Siezen, R. J. & Leunissen, J. A. M. Subtilases: The superfamily of subtilisin-like serine proteases. *Protein Sci.* **6**, 501–523 (1997).
- Kim, P. S. & Arvan, P. Endocrinopathies in the family of endoplasmic reticulum (ER) storage diseases: disorders of protein trafficking and the role of ER molecular chaperones. *Endocr. Rev.* **19**, 173–202 (1998).
- Hendershot, L. M. The ER chaperone BiP is a master regulator of ER function. *Mt. Sinai J. Med.* **71**, 289–297 (2004).
- Gething, M. J. Role and regulation of the ER chaperone BiP. *Semin. Cell Dev. Biol.* **10**, 465–472 (1999).
- Hamman, B. D., Hendershot, L. M. & Johnson, A. E. BiP maintains the permeability barrier of the ER membrane by sealing the luminal end of the translocon pore before and early in translocation. *Cell* **92**, 747–758 (1998).
- Lee, A. S. The ER chaperone and signaling regulator GRP78/BiP as a monitor of endoplasmic reticulum stress. *Methods* **35**, 373–381 (2005).
- Rao, R. V., Ellerby, H. M. & Bredesen, D. E. Coupling endoplasmic reticulum stress to the cell death program. *Cell Death Differ.* **11**, 372–380 (2004).
- Rao, R. V. *et al.* Coupling endoplasmic reticulum stress to the cell death program: role of the ER chaperone GRP78. *FEBS Lett.* **514**, 122–128 (2002).
- Jiang, J., Prasad, K., Lafer, E. M. & Sousa, R. Structural basis of interdomain communication in the Hsc70 chaperone. *Mol. Cell* **20**, 513–524 (2005).
- Le Bonniec, B. F., Guinto, E. R. & Esmon, C. T. Interaction of thrombin des-ETW with antithrombin III, the Kunitz inhibitors, thrombomodulin and protein C. Structural link between the autolysis loop and the Tyr-Pro-Trp insertion of thrombin. *J. Biol. Chem.* **267**, 19341–19348 (1992).
- Lencer, W. I. & Tsai, B. The intracellular voyage of cholera toxin: going retro. *Trends Biochem. Sci.* **28**, 639–645 (2003).
- Yu, M. & Haslam, D. B. Shiga toxin is transported from the endoplasmic reticulum following interaction with the luminal chaperone HEDJ/ERdj3. *Infect. Immun.* **73**, 2524–2532 (2005).
- Macario, A. J. L. & Conway de Macario, E. Sick chaperones, cellular stress and disease. *N. Engl. J. Med.* **353**, 1489–1501 (2005).
- Talbot, U. M., Paton, J. C. & Paton, A. W. Protective immunization of mice with an active-site mutant of subtilase cytotoxin of Shiga toxin-producing *Escherichia coli*. *Infect. Immun.* **73**, 4432–4436 (2005).
- Laemmli, U. K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680–685 (1970).
- Towbin, H., Staehelin, T. & Gordon, J. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc. Natl Acad. Sci. USA* **76**, 4350–4354 (1979).
- Maniatis, T., Fritsch, E. F. & Sambrook, J. Molecular cloning: a laboratory manual. (Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 1982).
- Kozutsumi, Y. *et al.* Identification of immunoglobulin heavy chain binding protein as glucose-regulated protein 78 on the basis of amino acid sequence, immunological cross-reactivity, and functional activity. *J. Cell Sci. Suppl.* **11**, 115–137 (1989).
- Yanisch-Perron, C., Vieira, J. & Messing, J. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13 mp18 and pUC19 vectors. *Gene* **33**, 103–119 (1985).
- Hobbs, S., Jitrapakdee, S. & Wallace, J. C. Development of a bicistronic vector driven by the human polypeptide chain elongation factor 1α promoter for creation of stable mammalian cell lines that express very high levels of recombinant proteins. *Biochem. Biophys. Res. Commun.* **252**, 368–372 (1998).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank M.-J. Gething and L. Helfenbaum for materials and advice, J. Wallace for discussions and L. Zaker-Tabrizi for technical assistance. This research was supported by a Program Grant from the Australian National Health and Medical Research Council (NHMRC) (to A.W.P. and J.C.P.). J.R. is supported by an Australian Research Council Professorial and Federation Fellowship and T.B. by a NHMRC Peter Doherty Fellowship. C.M.T. is supported by grants from the National Institutes of Health, USA.

Author Contributions A.W.P. designed, performed and interpreted experiments, and contributed to writing the manuscript. T.B., J.C.W., M.C.J.W. and J.R. crystallized SubA, solved the structure and contributed to manuscript preparation. C.M.T. contributed to experimental design and interpretation and writing of the manuscript. U.M.T. performed experiments. J.C.P. contributed to design and interpretation of experiments, project management and writing of the manuscript.

Author Information The coordinates and structure factors for SubA have been deposited in the Protein Data Bank and assigned the deposition code 2iy9. Reprints and permissions information is available at www.nature.com/reprints. The authors declare that they have no competing financial interests. Correspondence and requests for materials should be addressed to A.W.P. (adrianne.paton@adelaide.edu.au) or J.C.P. (james.paton@adelaide.edu.au).

Electron acceleration from contracting magnetic islands during reconnection

J. F. Drake¹, M. Swisdak², H. Che¹ & M. A. Shay³

A long-standing problem in the study of space and astrophysical plasmas is to explain the production of energetic electrons as magnetic fields 'reconnect' and release energy. In the Earth's magnetosphere, electron energies reach hundreds of thousands of electron volts (refs 1–3), whereas the typical electron energies associated with large-scale reconnection-driven flows are just a few electron volts. Recent observations further suggest that these energetic particles are produced in the region where the magnetic field reconnects⁴. In solar flares, upwards of 50 per cent of the energy released can appear as energetic electrons^{5,6}. Here we show that electrons gain kinetic energy by reflecting from the ends of the contracting 'magnetic islands' that form as reconnection proceeds. The mechanism is analogous to the increase of energy of a ball reflecting between two converging walls—the ball gains energy with each bounce. The repetitive interaction of electrons with many islands allows large numbers to be efficiently accelerated to high energy. The back pressure of the energetic electrons throttles reconnection so that the electron energy gain is a large fraction of the released magnetic energy. The resultant energy spectra of electrons take the form of power laws with spectral indices that match the magnetospheric observations.

The narrow current layers that form at the X-line during particle-in-cell simulations of magnetic reconnection spawn secondary magnetic islands⁷ (Fig. 1a, b), suggesting that the interaction of multiple magnetic islands is intrinsic to the dynamics of reconnection and associated particle acceleration. Electron acceleration during reconnection and island formation is reflected in the strong increase in the electron temperature parallel to the local magnetic field, $T_{e\parallel}$, in the simulation data (Fig. 1c). While some of these energetic electrons are produced as a result of acceleration by parallel electric fields near the magnetic separatrices and subsequent injection into the islands, electrons within the islands continue to gain energy. The energy gain, however, is not the result of a parallel electric field, which is very small within magnetic islands^{8,9}. The outflows from the magnetic X-lines cause the ends of the islands to move inward at the Alfvén speed. Particles circulating within the magnetic islands gain energy in a classic Fermi manner as they reflect from the moving ends of the islands (Fig. 2a, b). At the same time, the particles slowly drift outwards until they cross the separatrix and leave the island. The energy in the parallel motion is scattered into the perpendicular motion during the separatrix crossing (Fig. 2c)¹⁰. A similar Fermi process has been studied during Earthward-directed flows in the near-Earth region of the magnetotail¹¹, and is evident in the data of test particles in multi-island magnetohydrodynamic models^{12,13}.

The rate of gain of the parallel energy $\varepsilon_{\parallel}$ during reflection from the ends of magnetic islands contracting with a velocity u_x can be

calculated analytically (Fig. 2),

$$\frac{d\varepsilon_{\parallel}}{dt} = -2\varepsilon_{\parallel} \frac{u_x B_x^2}{\delta_x B^2} \quad (1)$$

where $2\delta_x$ is the length of the island, and B_x and B are the reconnecting and total magnetic fields. Equation (1) is independent of mass so that the rate of energy gain is the same for electrons and ions. However, the equation is only valid for particles with large parallel velocity $v_{\parallel} \gg u_x \approx c_{Ax}$, with c_{Ax} the Alfvén speed, so only the motion of very energetic ions is described by equation (1). Equation (1) can be integrated to obtain the invariant action ($\varepsilon_{\parallel}^{1/2} \delta_x$ for small B_x). Thus, the energy gain of the particles in a given island is limited by the change in the island geometry. To reach very high energies, particles must therefore interact with many islands. The trajectory of the particle in Fig. 2 suggests that this is a likely scenario.

The efficiency of electron heating by the Fermi mechanism is confirmed in particle simulations of the contraction of squashed flux bubbles (Fig. 3). An important question is how the back pressure from energetic particles limits island contraction. Using equation (1), the increase in the particle energy can be combined with the released magnetic energy to obtain an expression for the total energy change dW during an increment in the island length $d\delta_x$:

$$dW = 2 \frac{d\delta_x}{\delta_x} \frac{B_x^2}{8\pi} \left(1 - \frac{8\pi \bar{\varepsilon}_{\parallel}}{B^2} \right) \quad (2)$$

where the bar over $\varepsilon_{\parallel}$ denotes an average over the distribution of particle velocities. For negative $d\delta_x$, magnetic energy is released while electrons increase their energy and therefore inhibit contraction. When the back pressure of the energetic particles approaches the local magnetic energy density, island contraction ceases and halts the transfer of energy from the magnetic field into the energetic particles (Fig. 3c). Fermi acceleration during island contraction is therefore self-limiting and links the energy gain of particles to the released magnetic energy.

In a three-dimensional system, unless the ambient guide field is close to zero, magnetic islands are not constrained to form a single chain along the symmetry line of a current layer but can form at other locations¹⁴. Islands therefore should be volume filling in a region around a large-scale X-line (Fig. 4a). In such a picture, islands form, grow and contract. After contraction, the magnetic energy having been released, the island merges with a neighbour (lower current layer in Fig. 1b), completing its life cycle. Particles gain energy through Fermi acceleration in a sequence of islands. The shocks associated with many small islands may also accelerate particles¹⁵. As it is not possible to simulate a large-scale system in a three-dimensional kinetic model, we develop a transport model of particle heating. We average the energy gain equation over many islands (Fig. 4a), taking into account the scattering of energy from the parallel to the

¹University of Maryland, College Park, Maryland 20742, USA. ²Plasma Physics Division, Naval Research Laboratory, Washington DC 20375, USA. ³University of Delaware, Newark, Delaware 19716, USA.

perpendicular direction, and construct an equation for the omnidirectional distribution function $F(x, y, v) \equiv v^2 f(x, y, v)$:

$$\nabla \cdot \mathbf{u}F - \nabla \cdot \mathbf{D} \cdot \nabla F = -\frac{A}{3} \left(1 - \frac{8\pi\bar{\varepsilon}}{3B^2} \right)^{\frac{1}{2}} \frac{dC_{Ax}}{dy} \frac{\partial}{\partial v} vF \quad (3)$$

where $f(x, y, v)$ is the particle phase space density, $\mathbf{u}$ is the local plasma velocity, $\mathbf{D}$ is the diffusion rate¹⁶ and $\bar{\varepsilon} = 3\varepsilon_{\parallel}$. This equation describes the balance between the Fermi drive and the energy loss associated with convection and diffusion, including the back pressure from energetic particles. It is similar to that describing particle acceleration in shocks¹⁷. We solve this equation in the flow geometry shown in Fig. 4a, where the half-widths of the region of contracting islands are given by Δ_y and Δ_x in the inflow and outflow directions, respectively. Within the region of overlapping islands, convection is the dominant loss mechanism and the energy spectrum can be calculated analytically (see Supplementary Information):

$$F(v) = \frac{2\sigma - 1}{v^{2\sigma}} \int_0^v dq q^{2\sigma-1} F_{\text{in}}, \quad \sigma = \frac{1}{2} \left(\frac{1}{\hat{A}} + 1 \right) \quad (4)$$

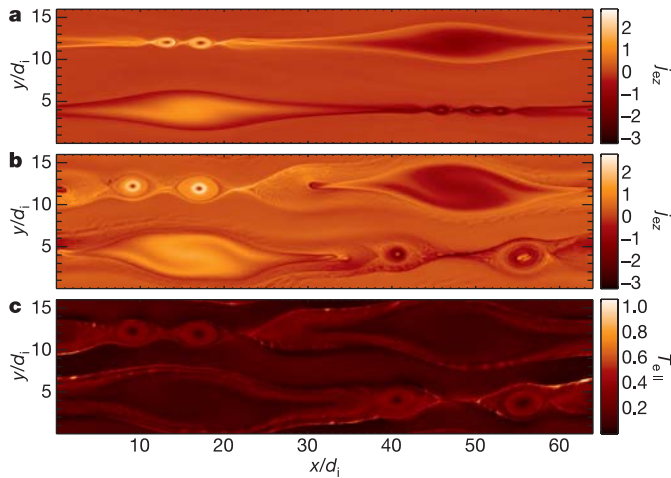


Figure 1 | Computer simulations of island formation and electron acceleration during magnetic reconnection. Particle-in-cell simulations using the p3d code²² are performed in doubly periodic two-dimensional geometry starting with two Harris current sheets with a peak density of n_0 superimposed on an ambient population of uniform density ($0.2n_0$). The reconnection magnetic field is:

$$B_x/B_0 = \tanh[(y - L_y/4)/w_0] - \tanh[(y - 3L_y/4)/w_0] - 1$$

where B_0 is the asymptotic magnetic field, $w_0 = 0.5d_i$, $L_x = 64d_i$ and $L_y = 16d_i$ are the half-width of the initial current sheets and the box size in the x and y directions. The electron and ion temperatures, respectively $T_e/m_i c_A^2 = 1/12$ and $T_i/m_i c_A^2 = 5/12$, are initially uniform as is the initial out-of-plane ‘guide’ field $B_z/B_0 = 1.0$. The ion inertial length is given by $d_i = c/\omega_{pi}$ with $\omega_{pi} = (4\pi n_0 e^2/m_i)^{1/2}$ and the Alfvén speed is given by $c_A = B_0/(4\pi m_i n_0)^{1/2}$. The electron mass m_e is taken to be $0.01m_i$ and the velocity of light $c = 20c_A$. The spatial grid consists of $4,096 \times 1,024$ cells with 100 particles per cell in the ambient background. The electron out-of-plane current j_{ez} is shown at two times: $t = 14.0\Omega_{ci}^{-1}$ in **a** and $t = 20.0\Omega_{ci}^{-1}$ in **b**, where $\Omega_{ci} = eB_0/m_i c$ is the ion cyclotron frequency. The spontaneous formation and growth of secondary magnetic islands is evident. The repeated breakup of X-line current layers is typical of guide-field reconnection where narrow current layers promote secondary island formation⁷. The electron temperature parallel to the local magnetic field, $T_{e\parallel}$, is shown at $t = 20.0\Omega_{ci}^{-1}$ in **c**. Seen is intense heating around the rims of the islands, which results from the acceleration of the electrons by parallel electric fields near the magnetic separatrices^{8,9}, and heating within the magnetic islands. The localization of the parallel electric field to the vicinity of the separatrix reduces its importance as an electron accelerator. The Fermi mechanism dominates when $v_{\parallel}$ exceeds the electron Alfvén speed $c_{Aex} = B_x/\sqrt{4\pi n m_e}$, which corresponds to an energy of 10 keV in the Earth’s magnetotail.

where $\hat{A} = A\Delta_x/(3\Delta_y)$ is the normalized Fermi drive, F_{in} is the upstream value of F and we have ignored the back pressure from the energetic particles. The distribution function at high energy takes the form of a power law with an index that depends on the mean aspect ratio of the individual magnetic islands and the island region through $\hat{A}$. Numerical solutions to equation (3) for F confirm the power-law behaviour (Fig. 4b, c). The aspect ratio of the magnetic islands, and therefore $\hat{A}$ and σ , remain uncertain. For $\hat{A} > 0.5$ or $\sigma < 1.5$, the energy integral of the energetic particles $\bar{\varepsilon}$ diverges unless the back pressure of energetic particles is included. Retaining the back pressure, the energy content of electrons rises until $\bar{\varepsilon} \approx B^2/8\pi$. The back pressure throttles the Fermi drive and the spectral index of the energetic particles can be characterized by the upstream value of the electron thermal (p_{e0}) to magnetic pressure ratio $\beta_{e0} = 8\pi p_{e0}/B^2$, independent of $\hat{A}$ (Fig. 4c).

The predictions of the model can be compared with several key observations in the magnetotail. The isotropic spectrum observed above an energy threshold in the Wind satellite observations⁴ results from scattering as particles pass close to X-lines. Particle energies well in excess of the potential drop across the tail^{1,2} with dawn–

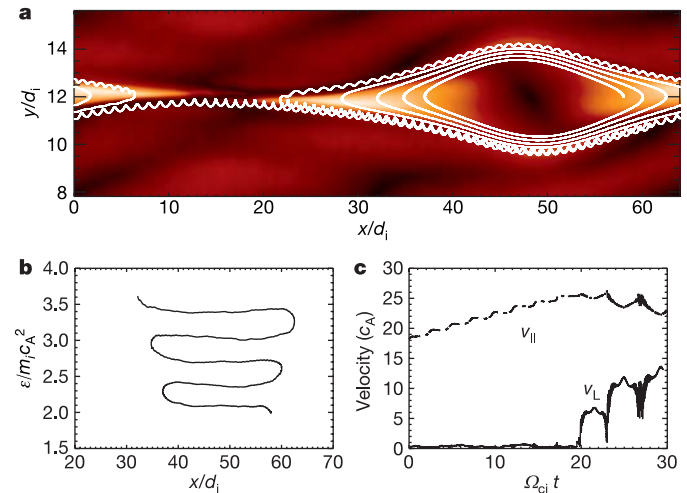


Figure 2 | Test particle orbits and energy gain of Fermi accelerated electrons.

The orbits are computed from the fields of the simulation in Fig. 1 at a time $t = 10.8\Omega_{ci}^{-1}$, just before the formation of secondary islands. **a**, The orbit of a particle started at the midplane on the right side of the upper island $x, y = 58.0d_i, 12.0d_i$ with an initial velocity given by the local $\mathbf{E} \times \mathbf{B}$ velocity plus a parallel velocity $v_{\parallel}$ of $10.8c_A$ shown on the background of **E**. The particle follows field lines and slowly drifts outward. **b**, The particle energy ε as a function of its x position. The particle gains energy as it reflects from the ends of the islands, which are moving inwards at the Alfvén speed. The energy gain therefore results from a classical Fermi reflection. Because the velocity of energetic electrons greatly exceeds the Alfvén speed, many reflections are required for electrons to reach high energy. Also evident in **a** is the sudden change in the orbit as the island approaches the separatrix—the gyration radius of the particle abruptly increases as the particle encounters the sharp kink in the magnetic field line just downstream from the X-line at $x, y = 16d_i, 12d_i$. **c**, The parallel velocity, $v_{\parallel}$, which increases in time until $t = 20\Omega_{ci}^{-1}$ when the local gyration velocity v_L abruptly increases. The separatrix crossing therefore scatters energy from the parallel into the perpendicular motion¹⁰. The energy gain during the reflection from the island ends in **b** can be calculated in a simple model in which B_y and B_z are constant and $B_x(y)$ increases away from the centre of the current layer. The reconnection field E_z and an in-plane electric field $E_y = -E_z B_z/B_y$ are chosen so that $\mathbf{E}_{\parallel} = \mathbf{b} \cdot \mathbf{E} = 0$. For electrons with $v_{\parallel} \gg v_L$, the change in the parallel velocity results from the curvature drift in the direction of the electric field, $dv_{\parallel}/dt = cv_{\parallel} \mathbf{E} \cdot \mathbf{b} \times \kappa/B$ where $\kappa = \mathbf{b} \cdot \nabla \mathbf{b}$. This equation can be integrated to obtain the increment in the parallel velocity $\delta v_{\parallel} = -2u_x B_x/B$ due to its reflection, where $u_x = -cE_z/B_y$ is the local velocity of the end of the island and B_x is given by its asymptotic value. The resulting rate of energy gain is given in equation (1).

dusk asymmetries that are weak during storm times^{3,18} require an acceleration mechanism that can diffuse particles downward and duskward while still gaining energy. In the Fermi model, the particle energy gain is independent of the sign of $v_{\parallel}$ and therefore independent of the direction of motion across the tail. Thus, there is no intrinsic limit on the energy gain. Distributions of high-energy particles in the magnetosphere are typically power laws, consistent with Fig. 4c. Remarkably, the model yields a spectral index for energetic particles of 3.7 for the Wind parameters, which is close to the measured value of 3.8 (ref. 4).

An important conclusion from solar satellite observations is that up to half of the energy released during solar flares is transferred to electrons. In the data presented in Fig. 4, the electrons receive the bulk of the released magnetic energy in the core of the multi-island region. In the low β_{e0} regime of the corona, the spectral index of energetic electrons approaches a lower limit of 1.5, a spectrum that is hard but is occasionally observed¹⁹. As essentially all of the electrons entering the multi-island region through the area $S = \Delta_x \times \Delta_z$ in Fig. 4a undergo significant acceleration, rates of energetic electron production of 10^{36} s^{-1} as inferred from observations²⁰ are possible but only if S is macroscopic ($n = 10^9 \text{ cm}^{-3}$, $u_y = 0.1c_{Ax}$, $B = 100 \text{ G}$ and $S = 10^{19} \text{ cm}^2$, yields $\dot{N} = nu_y S \sim 10^{36} \text{ s}^{-1}$). Some solar observations do reveal island-like structures in macroscale current layers²¹.

Data from the present fleet of magnetospheric satellites should be able to validate key features of the present model. Secondary

magnetic islands in the vicinity of the dissipation region are required for the production of very energetic electrons and should be measurable in the existing data sets. Secondary islands and energetic electrons should not be produced unless the ambient electron pressure is small ($\beta_{e0} = 8\pi p_{e0}/B^2 \ll 1$).

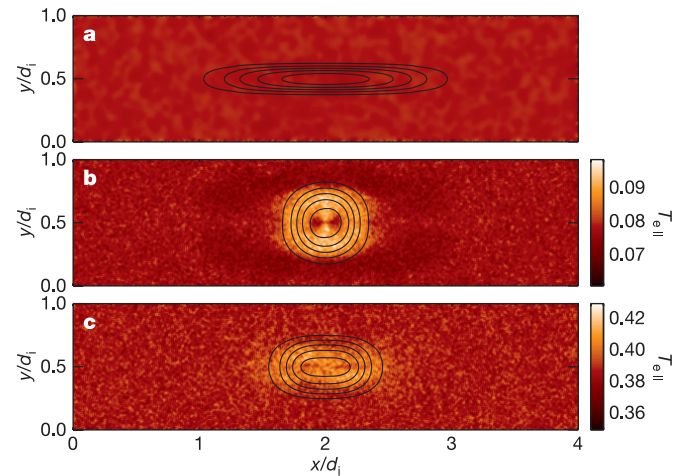


Figure 3 | Electron Fermi acceleration in squashed flux bubbles. Shown are the results of two-dimensional simulations of electron acceleration in isolated, squashed flux bubbles, where Fermi acceleration can be studied without the interference from the parallel electric fields associated with magnetic reconnection. The release of magnetic energy as the bubbles contract and become round is essentially the same as in the contraction of magnetic islands during magnetic reconnection. **a**, The initial magnetic field lines in a squashed magnetic bubble superimposed over the uniform, isotropic electron temperature $T_{e0} = 0.1$. The bubble is initially in force balance in the y direction (the variation in the density enabling the plasma pressure to balance the magnetic pressure) but not in the x direction. Initially $\beta_{e0} = 8\pi n_0 T_{e0}/B_0^2 = 0.27$, where n_0 and B_0 are the maximum values of the density and magnetic field, and B_z is zero. Realistic values of the electron mass ($m_i/1,836$) and velocity of light ($c = 100c_A$) are required so the electrons can undergo many Fermi reflections during the contraction of the bubble. **b**, The parallel electron temperature $T_{e||}$ and magnetic field lines at late time. Note the increase in the parallel temperature within the bubble and the near circular shape of the final magnetic field lines. The change in the perpendicular temperature is small. Sixty percent of the released magnetic energy is transferred to electrons. **c**, $T_{e||}$ and the magnetic field lines at late time shown from a simulation that is identical to that shown in **a** and **b** but with a larger initial electron pressure ($\beta_{e0} = 1.1$). In this case, the back pressure from the accelerated electrons prevents the full contraction of the bubble, consistent with the discussion following equation (2).

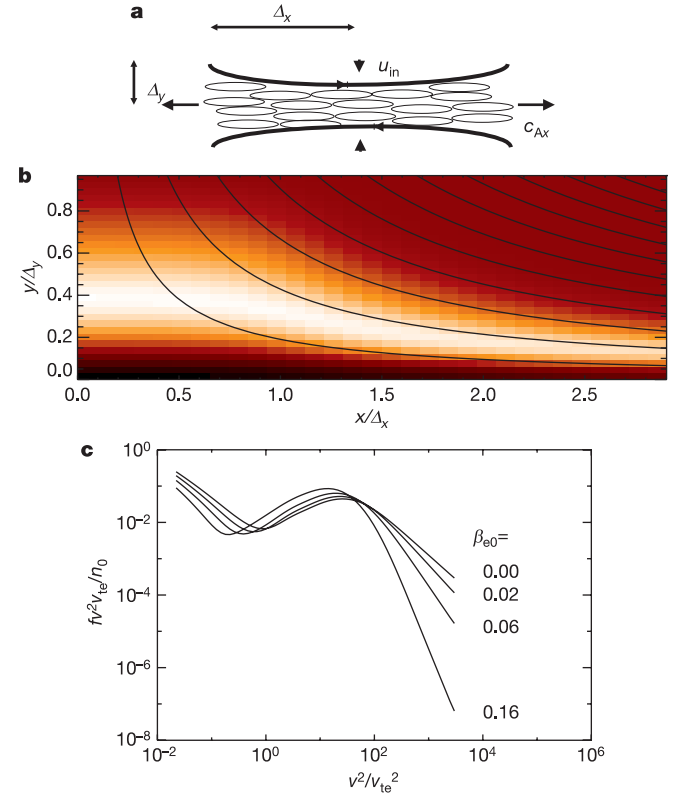


Figure 4 | Particle acceleration in a multi-island reconnection geometry.

a, Diagram showing volume filling islands expected around the reversal region, where owing to incompressibility the inflow velocity is given by $u_{in} = c_{Ax} \Delta_y / \Delta_x$. To calculate electron acceleration, we average the energy gain in equation (1) over many islands, including the scattering of energy from the parallel to the perpendicular direction:

$$\frac{d\varepsilon}{dt} = \frac{2\varepsilon}{3} \left\langle \frac{B_{xi}^2 c_{Axi}}{B_i^2 \delta_{xi}} \right\rangle = 2 \frac{A\varepsilon}{3} \frac{dc_{Ax}}{dy} \quad A \equiv \left\langle \frac{B_{xi}^2 \delta_{yi}}{B_i^2 \delta_{xi}} \right\rangle$$

where the contraction velocity of each island is the upstream Alfvén speed c_{Axi} and δ_{yi} is the island width. The subscript “i” indicates that the value can vary among the different islands. The resulting equation for the omnidirectional distribution function $F(x,y,v)$ is solved numerically. One quadrant ($x,y > 0$) of $F(x,y,1.95v_{te})$ is shown in **b** along with the streamlines of the flow for $\hat{A} = 0.6$ and $\beta_{e0} = 0.02$, where v_{te} is the initial thermal velocity of the upstream Maxwellian distribution. The white band results from the heating of inflowing low energy electrons and the dark band from the depletion of these particles as they move to even higher energies. **c**, Energy spectra for $\hat{A} = 0.6$ and several values of β_{e0} at $x,y = 0,0$. The spectra are power laws at high energy with spectral indices σ of 1.3, 1.7, 2.2 and 3.7 for $\beta_{e0} = 0.0, 0.02, 0.06$ and 0.16 , respectively. The dip results from the depletion of particles undergoing acceleration to higher energy. For $\hat{A} = 0.6$, the energy integrals of the spectra diverge in the absence of the back pressure from energetic particles (the $\beta_{e0} = 0.0$ spectrum in **c** matches the $\sigma = 1.33$ prediction from equation (4)). For any finite β_{e0} the back pressure from the energetic particles reduces the drive to regularize the energy spectrum. The limiting spectral index for low β_{e0} is 1.5. Higher values of β_{e0} yield softer spectra with $\sigma = 3.7$ for the case $\beta_{e0} = 0.16$, close to the measured value of 3.8 at $\beta_{e0} = 0.16$ from the Wind energetic particle observations⁴. In the strongly driven regime ($\hat{A} > 0.5$), the power-law indices are nearly independent of $\hat{A}$. The ambient value of β_{e0} is therefore the dominant control parameter for the spectra of energetic electrons during magnetic reconnection.

Received 30 August 2005; accepted 19 July 2006.

1. Terasawa, T. & Nishida, A. Simultaneous observations of relativistic electron bursts and neutral-line signatures in the magnetotail. *Planet. Space Sci.* **24**, 855–856 (1976).
2. Baker, D. N. & Stone, E. C. Energetic electron anisotropies in the magnetotail: identification of open and closed field lines. *Geophys. Res. Lett.* **3**, 557–560 (1976).
3. Meng, C.-I., Lui, A. T. Y., Krimigis, S. M. & Ismail, S. Spatial distribution of energetic particles in the distant magnetotail. *J. Geophys. Res.* **86**, 5682–5700 (1981).
4. Øieroset, M., Lin, R. P., Phan, T. D., Larson, D. E. & Bale, S. D. Evidence for electron acceleration up to 300 keV in the magnetic reconnection diffusion region in the earth's magnetotail. *Phys. Rev. Lett.* **89**, 195001 (2002).
5. Lin, R. P. & Hudson, H. S. 10–100 keV electron acceleration and emission from solar flares. *Sol. Phys.* **17**, 412–435 (1971).
6. Lin, R. P. *et al.* RHESSI observations of particle acceleration and energy release in an intense solar gamma-ray line flare. *Astrophys. J.* **595**, L69–L76 (2003).
7. Drake, J. F., Swisdak, M., Schoeffler, K. M., Rogers, B. N. & Kobayashi, S. Formation of secondary islands during magnetic reconnection. *Geophys. Res. Lett.* **33**, L13105, doi:10.1029/2006GL025957 (2006).
8. Pritchett, P. L. & Coroniti, F. V. Three-dimensional collisionless magnetic reconnection in the presence of a guide field. *J. Geophys. Res.* **109**, A01220, doi:10.1029/2003JA009999 (2004).
9. Drake, J. F., Shay, M. A., Thongthai, W. & Swisdak, M. Production of energetic electrons during magnetic reconnection. *Phys. Rev. Lett.* **94**, 095001 (2005).
10. Buechner, J. & Zelenyi, L. M. Regular and chaotic particle motion in sheared magnetic field reversals. *Adv. Space Res.* **11**, 177–182 (1991).
11. Birn, J., Thomsen, M. F. & Hesse, M. Electron acceleration in the dynamic magnetotail: Test particle orbits in three-dimensional magnetohydrodynamic simulation fields. *Phys. Plasmas* **11**, 1825–1833 (2004).
12. Matthaeus, W. H., Ambrosiano, J. J. & Goldstein, M. J. Particle acceleration by turbulent magnetohydrodynamic reconnection. *Phys. Rev. Lett.* **53**, 1449–1452 (1984).
13. Kliem, B. Particle orbits, trapping and acceleration in a filamentary current sheet model. *Astrophys. J.* **90**, 719–727 (1994).
14. Galeev, A. A., Kuznetsova, M. M. & Zelenyi, L. M. Magnetopause stability threshold for patchy reconnection. *Space Sci. Rev.* **44**, 1–41 (1986).
15. Shibata, K. & Tanuma, S. Plasmoid-induced-reconnection and fractal reconnection. *Earth Planets Space* **53**, 473–482 (2001).
16. Jokipii, J. R. & Parker, E. N. Stochastic aspects of magnetic field lines of force with application to cosmic-ray propagation. *Astrophys. J.* **155**, 777–798 (1969).
17. Blandford, R. D. & Eichler, D. Particle acceleration at astrophysical shocks: a theory of cosmic ray origin. *Phys. Rep.* **154**, 1–75 (1987).
18. Imada, S., Hoshino, M. & Mukai, T. in *Frontiers in Magnetospheric Plasma Physics: Celebrating 10 Years of Geotail Operation* (eds Hoshino, M., Omura, Y. & Lanzerotti, L. J.) 34–39 (Cospar Colloquia Series Vol.16, Elsevier, Oxford, 2005).
19. Holman, G. D. Energetic electrons in solar flares as viewed in x-rays. *Adv. Space Res.* **35**, 1669–1674 (2005).
20. Miller, J. A. *et al.* Critical issues for understanding particle acceleration in impulsive solar flares. *J. Geophys. Res.* **102**, 14631–14659 (1997).
21. Lin, J. *et al.* Direct observations of the magnetic reconnection site of an eruption on 2003 November 18. *Astrophys. J.* **622**, 1251–1264 (2005).
22. Zeiler, A. *et al.* Three-dimensional particle simulations of collisionless magnetic reconnection. *J. Geophys. Res.* **107**, 1230, doi:10.1029/2001JA000287 (2002).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This work was supported by the NSF/DOE programme in plasma science, by NASA through the Supporting Research and Technology and the Sun-Earth Connections Theory programmes, through CMPD, a DOE FSC, and through CISM, an NSF STC. Computations were carried out in part at the National Energy Research Scientific Computing Center.

Author Contributions J.F.D., M.S. and M.A.S. identified the Fermi mechanism; J.F.D. carried out the particle simulations of reconnection, obtained the solutions of the transport equation and wrote the paper; and M.S. and H.C. carried out the bubble and test particle simulations, respectively. All of the authors discussed the results and commented on the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to J.F.D. (drake@plasma.umd.edu).

Quantum teleportation between light and matter

Jacob F. Sherson^{1,3}, Hanna Krauter¹, Rasmus K. Olsson¹, Brian Julsgaard¹, Klemens Hammerer², Ignacio Cirac² & Eugene S. Polzik¹

Quantum teleportation¹ is an important ingredient in distributed quantum networks², and can also serve as an elementary operation in quantum computers³. Teleportation was first demonstrated as a transfer of a quantum state of light onto another light beam^{4–6}; later developments used optical relays⁷ and demonstrated entanglement swapping for continuous variables⁸. The teleportation of a quantum state between two single material particles (trapped ions) has now also been achieved^{9,10}. Here we demonstrate teleportation between objects of a different nature—light and matter, which respectively represent ‘flying’ and ‘stationary’ media. A quantum state encoded in a light pulse is teleported onto a macroscopic object (an atomic ensemble containing 10^{12} caesium atoms). Deterministic teleportation is achieved for sets of coherent states with mean photon number (n) up to a few hundred. The fidelities are 0.58 ± 0.02 for $n = 20$ and 0.60 ± 0.02 for $n = 5$ —higher than any classical state transfer can possibly achieve¹¹. Besides being of fundamental interest, teleportation using a macroscopic atomic ensemble is relevant for the practical implementation of a quantum repeater². An important factor for the implementation of quantum networks is the teleportation distance between transmitter and receiver; this is 0.5 metres in the present experiment. As our experiment uses propagating light to achieve the entanglement of light and atoms required for teleportation, the present approach should be scalable to longer distances.

Quantum teleportation—a disembodied transfer of a quantum state with the help of distributed entanglement—was proposed in a seminal paper¹. The generic protocol of quantum teleportation begins with the creation of a pair of entangled objects which are shared by two parties, Alice and Bob. This step establishes a quantum link between them. Alice receives an object to be teleported and performs a joint measurement on this object and her entangled object (a Bell measurement). The result of this measurement is communicated via a classical communication channel to Bob, who uses it to perform local operations on his entangled object, thus completing the process of teleportation.

In our experiment, a pair of entangled objects is created by sending a strong ‘in’ pulse of light (shown on the left in Fig. 1) through an atomic sample at Bob’s location. As a result of the interaction between the light and the atoms, the transmitted ‘out’ light received by Alice’s and Bob’s atoms become entangled. On Alice’s site the entangled pulse is mixed with the pulse to be teleported on a 50/50 beamsplitter (BS in Fig. 1). A Bell measurement in the form of homodyne measurements of the optical fields in the two output ports of the BS is carried out and the results are transferred to Bob as classical photocurrents. Bob performs spin rotations on the atoms to complete the teleportation protocol. Finally, the state of the atoms is analysed to confirm that the teleportation has been successful.

The experiment follows a recent proposal for light-to-atoms teleportation¹² using multimode entanglement of light with an

atomic ensemble placed in a magnetic field. We describe teleportation in the language of dimensionless canonical variables¹³; this provides a common description for light and atoms, and allows for a complete tomographic characterization of the states.

The atomic object is a spin-polarized gas sample of approximately $N_{\text{at}} = 10^{12}$ caesium atoms in a $25 \times 25 \times 25$ mm paraffin-coated glass cell at around room temperature^{14–18} placed in a homogeneous magnetic field ($\mathbf{B}$). Atoms are initially prepared in a coherent spin state by a 4-ms circularly polarized optical pumping pulse propagating along the direction of the magnetic field, into the sublevel $F = 4$, $m_F = 4$ (Fig. 1) of the ground state with the collective ensemble angular momentum $\langle \hat{J}_x \rangle = J_x = 4N_{\text{at}}$, and the transverse projections with minimal quantum uncertainties, $\langle \delta J_y^2 \rangle = \langle \delta J_z^2 \rangle = \frac{1}{2}J_x$. Changing to the frame rotating at the Larmor frequency Ω and introducing the canonical variables for the collective transverse atomic spin components¹², we obtain $\hat{X}_A = \hat{J}_y^{\text{rot}} / \sqrt{J_x}$, $\hat{P}_A = \hat{J}_z^{\text{rot}} / \sqrt{J_x}$ which obey

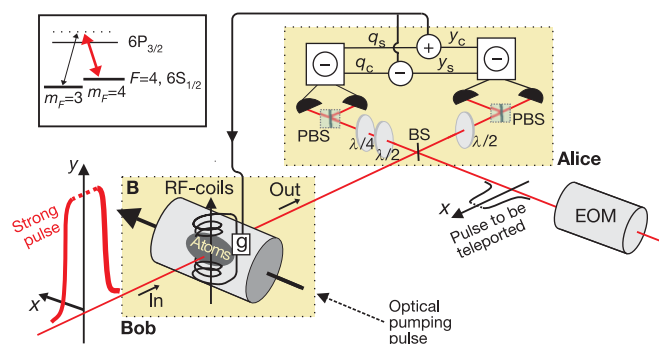


Figure 1 | Experimental set-up for teleportation of light onto an atomic ensemble. Atoms are initially optically pumped into $F = 4$, $m_F = 4$ state with a 4-ms pulse. A strong y-polarized 2-ms ‘in’ pulse of light is then sent through the atomic sample at Bob’s location and becomes entangled with the atoms (the pulse length is around 600 nm and is not shown to scale in the figure). The pulse travels 0.5 m to Alice’s location, where it is mixed on a beamsplitter (BS) with the object of teleportation—a few-photon coherent pulse of light—generated by the electro-optical modulator (EOM) synchronously with the strong pulse. In the two output ports of the BS, two polarization beamsplitters (PBS) split light onto two pairs of detectors which perform a polarization homodyne measurement (a Bell measurement). The results of these measurements are combined, processed electronically, as described in the text, and sent via a classical communication channel to Bob. There they are used to complete the teleportation onto atoms by shifting the atomic collective spin state with a pulse of a radio-frequency (RF) magnetic field of 0.2-ms duration. After a delay of 0.1 ms, a second strong pulse—the verifying pulse—is sent to read out the atomic state, in order to prove the successful teleportation. Inset, relevant atomic sublevels and light modes (not to scale). The frequency difference between a weak quantum field (black arrow) and the strong entangling field (thick red arrow) is equal to the Zeeman splitting of the ground state sublevels.

¹Niels Bohr Institute, Copenhagen University, Blegdamsvej 17, Copenhagen Ø, Denmark. ²Max Planck Institute for Quantum Optics, Hans-Kopfermann-Str. 1, Garching, D-85748, Germany. ³Department of Physics and Astronomy, University of Aarhus, Aarhus, 8000, Denmark.

the canonical commutation relation $[\hat{X}_A, \hat{P}_A] = i$ provided that $J_x \gg \sqrt{\langle \delta J_{y,z}^2 \rangle}, \langle J_{y,z} \rangle$. Here $\hat{X}_A$ and $\hat{P}_A$ are the recipient operators in the teleportation protocol.

The light to be teleported, and the ‘in’ and ‘out’ modes (Fig. 1), are described by single mode canonical operators^{6,12} $\hat{Y}, \hat{Q}$, and $\hat{y}^{\text{in}}, \hat{q}^{\text{in}}$ and $\hat{y}^{\text{out}}, \hat{q}^{\text{out}}$, respectively. These operators obeying $[\hat{Y}, \hat{Q}] = [\hat{y}, \hat{q}] = i$ are quantum analogues of the amplitude and phase of light in classical physics, or, more precisely, of the classical quadrature phase amplitudes y, q in the decomposition of the electric field of light with the frequency ω as $E \propto y \cos \omega t + q \sin \omega t$ (see Methods for exact definitions). Two non-commuting variables in quantum mechanics cannot be measured without distortion. The challenge of teleportation thus consists of a faithful transfer of these not simultaneously measurable operators, $\hat{Y}, \hat{Q}$, onto atomic operators $\hat{X}_A$ and $\hat{P}_A$. The Raman-type interaction (see Fig. 1 inset) couples the quantum $\omega + \Omega$ sideband of the ‘in’ field to the Zeeman sublevels separated by the frequency $\Omega = 322$ kHz. Therefore we introduce the $\cos \Omega t, \sin \Omega t$ components of the light operators $\hat{Y}_{c,s}, \hat{Q}_{c,s}$ and $\hat{y}_{c,s}, \hat{q}_{c,s}$ (see Methods). Canonical operators for the upper sideband mode $\hat{Y}, \hat{Q}$ can be expressed¹² via measurable $\sin(\Omega t)$ and $\cos(\Omega t)$ components, $\hat{Y}_s, \hat{Q}_s, \hat{Y}_c, \hat{Q}_c$, as $\hat{Y} = \frac{1}{\sqrt{2}}(\hat{Y}_s + \hat{Q}_c), \hat{Q} = -\frac{1}{\sqrt{2}}(\hat{Y}_c - \hat{Q}_s)$.

We first describe generation of entanglement between light and atoms. The ‘in’ strong pulse is y -polarized, hence its x -polarized mode $\hat{y}^{\text{in}}, \hat{q}^{\text{in}}$ is in a vacuum state. After interaction with atoms¹², the x -polarized ‘out’ mode operators $\hat{y}^{\text{out}}, \hat{q}^{\text{out}}$ are given by:

$$\begin{aligned} \hat{y}_c^{\text{out}} &= \left\{ \hat{y}_c^{\text{in}} + \frac{\kappa^2}{4} \hat{q}_s^{\text{in}} + \frac{\kappa^2}{4\sqrt{3}} v_s \right\} + \frac{\kappa}{\sqrt{2}} \hat{P}_A^{\text{in}}, & \hat{q}_{s,c}^{\text{out}} &= \hat{q}_{s,c}^{\text{in}} \\ \hat{y}_s^{\text{out}} &= \left\{ \hat{y}_s^{\text{in}} - \frac{\kappa^2}{4} \hat{q}_c^{\text{in}} - \frac{\kappa^2}{4\sqrt{3}} v_c \right\} - \frac{\kappa}{\sqrt{2}} \hat{X}_A^{\text{in}} \end{aligned} \quad (1)$$

The terms in curly brackets in the equations for $\hat{y}$ represent vacuum contributions coming from different orthogonal modes of the ‘in’ pulse where the canonical operators $v_{s,c}$ represent vacuum temporal higher order canonical modes¹². The terms containing $\hat{P}_A^{\text{in}}$ and $\hat{X}_A^{\text{in}}$ describe the imprint of the atomic state on the light via coherent forward scattering from the atomic ensemble, or, in other words, polarization rotation due to the Faraday effect^{4,15}. The atomic spin operators are transformed by the interaction with light as follows¹²:

$$\hat{X}_A^{\text{out}} = \hat{X}_A^{\text{in}} + \frac{\kappa}{\sqrt{2}} \hat{q}_c^{\text{in}}, \quad \hat{P}_A^{\text{out}} = \hat{P}_A^{\text{in}} + \frac{\kappa}{\sqrt{2}} \hat{q}_s^{\text{in}} \quad (2)$$

The second terms in equation (2) describe the imprint of the light state onto atoms via the dynamic Stark effect¹⁴.

The atoms–light entanglement described by equations (1) and (2) is very close¹², under our experimental conditions, to the Einstein–Podolsky–Rosen entanglement optimal for quantum teleportation. The atoms–light coupling constant $\kappa = a_1 \sqrt{N_{\text{ph}} N_{\text{at}}} \Gamma \sigma / \Delta \propto \alpha_0$

has been discussed in detail previously^{12,14–18}. (Here σ is the dipole cross-section¹⁸, a_1 is the vector polarizability¹⁸, $\Gamma = 2.6$ MHz is the natural linewidth (HWHM) of the transition, $N_{\text{ph}} = 4 \times 10^{13}$ is the number of the y -polarized photons in the strong pulse, $\Delta = 825$ MHz is the blue detuning of light from the atomic resonance, and $A = 4.8$ cm² is the cross-section of the atomic sample.) As in our previous experiments with the atoms–light quantum interface, strong coupling with the atomic ensemble is achieved in the region of a high resonant optical depth α_0 . In the experiment we choose a nearly optimal value¹² of $\kappa \approx 1$ by changing $\alpha_0 \propto N_{\text{at}}$ with the temperature of the vapour. Note that another condition for strong coherent coupling is a very high N_{ph} in the y -polarized mode.

At Alice’s location (Fig. 1), the ‘out’ pulse is mixed on BS with the object of teleportation—a few-photon x -polarized coherent pulse with frequency $\omega + \Omega$ generated by an electro-optical modulator (EOM). A Bell measurement of canonical variables^{6,12,13} is performed by two sets of polarization homodyne detectors in the two output ports of BS (Fig. 1). Homodyne detection followed by the normalization to the vacuum (shot) noise of light⁶ is a standard method for measuring canonical variables of light. In our experiment, the strong y -polarized pulse, besides driving the entangling interaction, also plays the role of a local oscillator for the homodyne detection. The variables in phase with the strong pulse $\hat{y}_{c,s} = \frac{1}{\sqrt{2}}(\hat{y}_{c,s}^{\text{out}} + \hat{Y}_{c,s})$ are measured via a measurement of the Stokes parameter $\hat{S}_2$ in one output of BS, whereas the out-of-phase components $\hat{q}_{c,s} = \frac{1}{\sqrt{2}}(\hat{q}_{c,s}^{\text{out}} - \hat{Q}_{c,s})$ are measured via the Stokes parameter $\hat{S}_3$ in the other arm (see Methods). The $\sin(\Omega t)$ and $\cos(\Omega t)$ components are measured by processing photocurrents with lock-in amplifiers. The Bell measurement of operators $\hat{y}_{c,s}$ and $\hat{q}_{c,s}$ yields four results, $y_{c,s}$ and $q_{c,s}$. Operationally, these values are properly normalized integrals of corresponding photocurrents over the pulse duration (see Methods). As shown in Fig. 1, the photocurrents are combined to yield two feedback signals proportional to $y_s - q_c$ and $y_c + q_s$ which are sent from Alice to Bob. Auxiliary magnetic field pulses^{14,17} with frequency Ω and amplitudes proportional to the feedback signals are applied to the atoms, so that the collective atomic spin variables at Bob’s site are shifted to become:

$$\begin{aligned} \hat{X}_A^{\text{tele}} &= \hat{X}_A^{\text{out}} + g_x(y_s - q_c) = \hat{X}_A^{\text{out}} + \frac{1}{\sqrt{2}} g_x (\hat{y}_s^{\text{out}} - \hat{q}_c^{\text{out}}) + g_x \hat{Y} \\ \hat{P}_A^{\text{tele}} &= \hat{P}_A^{\text{out}} - g_p(y_c + q_s) = \hat{P}_A^{\text{out}} - \frac{1}{\sqrt{2}} g_p (\hat{y}_c^{\text{out}} + \hat{q}_s^{\text{out}}) + g_p \hat{Q} \end{aligned} \quad (3)$$

where $g_{x,p}$ are the feedback gains. This step completes the teleportation protocol, as the light operators $\hat{Y}, \hat{Q}$ are now transferred onto atomic operators $\hat{X}_A^{\text{tele}}, \hat{P}_A^{\text{tele}}$, and all other terms in equation (3) can be made small with a suitable choice of κ and g .

To prove that we have performed the quantum teleportation, we determine the fidelity of the teleportation. Towards this end, we send a second—verifying—strong pulse of y -polarized light through the

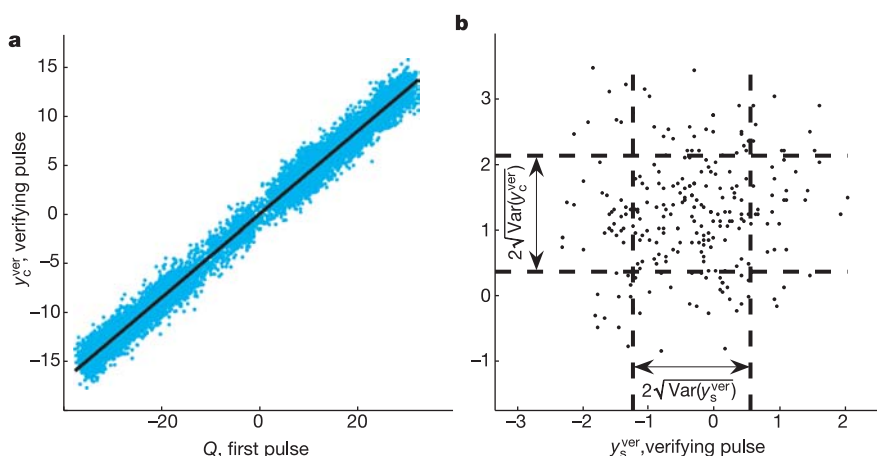


Figure 2 | Raw experimental data for a series of teleportation runs. **a**, Calibration of the teleportation feedback gain. Verifying pulse canonical variable y_c^{ver} versus the input pulse canonical variable Q for 10,000 teleportation runs. All dimensionless canonical variables are normalized so that their variance for a vacuum state is 1/2. The coherent input state used in the plot has a mean photon number of $\bar{n} \approx 500$, and is slowly modulated in phase during this measurement. The straight line fit is used for calibration of the feedback gain (see comments in the text). **b**, An example of data from which the atomic state variances after the teleportation are determined. Two canonical variables of the verifying pulse, y_s^{ver} and q_s^{ver} , are plotted for an input state with $\bar{n} = 5$ and a fixed phase. The dashed lines indicate twice the standard deviation intervals $2\sqrt{\text{Var}(y_{c,s}^{\text{ver}})}$ which are used to determine the atomic state variances as discussed in the text.

atomic ensemble after the teleportation is completed. From this measurement we reconstruct the atomic operators $\hat{X}_A^{\text{tele}}$ and $\hat{P}_A^{\text{tele}}$. The fidelity is the overlap of the input state and the teleported state averaged over the input state distribution^{12,14,17}. The classical benchmark fidelity which has to be exceeded in order to claim the success of quantum teleportation is known¹¹ for a gaussian distribution of coherent states with the width corresponding to the mean photon number $\langle n \rangle$ centred at zero. The experimental fidelity for such distribution can be found as^{6,18}:

$$F_n = \frac{2}{\sqrt{(2\langle n \rangle(1 - g_X)^2 + 1 + 2\sigma_X^2)(2\langle n \rangle(1 - g_P)^2 + 1 + 2\sigma_P^2)}}$$

The gains are defined from the mean values of atomic and light operators: $\bar{X}_A^{\text{tele}} = g_X \bar{Y}$, $\bar{P}_A^{\text{tele}} = g_P \bar{Q}$. σ_X^2, σ_P^2 are the variances for the final gaussian state of the atoms.

The mean values for the input light operators are determined from the results of the Bell measurement: $\bar{y}_s - \bar{q}_c = \bar{Y}$ and $\bar{y}_c + \bar{q}_s = \bar{Q}$. The mean values and the variances of the atomic operators are determined from the verifying pulse measurements. Using equations (1) and (3) and the input–output beamsplitter relations¹², we can link the measurement of the verifying pulse on the S_2 detector to the atomic mean values: $\bar{y}_c^{\text{ver}} = \frac{\kappa}{2} \bar{P}_A^{\text{tele}} = \frac{g_P \kappa}{2} \bar{Q}$, $\bar{y}_s^{\text{ver}} = \frac{\kappa}{2} \bar{X}_A^{\text{tele}} = \frac{g_X \kappa}{2} \bar{Y}$. Using these expressions, we can calibrate $g_{X,P}$, as shown in Fig. 2a where $\bar{y}_c^{\text{ver}}$ is plotted as a function of $\bar{Q}$, as the value of $\kappa = 0.93$ is determined independently from the projection noise measurement (see Methods). From the linear fit to this distribution we find g_P , which can then be tuned to a desired value electronically. Results plotted in Fig. 2a along with similar results for the other operator $\bar{y}_s(\bar{Y})$ present the proof of the successful classical transfer of the mean values of the quantum mechanical operators $\bar{Y}, \bar{Q}$ of light onto atomic operators.

To verify the success of the quantum teleportation, we have to determine the variances of the two atomic operators which now contain the teleported input light operators. Figure 2b shows an example of results $\bar{y}_c^{\text{ver}}, \bar{y}_s^{\text{ver}}$ for 250 teleportation runs for a fixed input state. Making use of equation (1) and the beamsplitter relations, we can directly find the atomic state variances from $\text{Var}\{\hat{y}_{s(c)}\}$ of such distribution as $\sigma_{X(P)}^2 = \frac{4}{\kappa^2} [\text{Var}\{\hat{y}_{s(c)}\} - \frac{\kappa^4}{48} - \frac{1}{2}]$. The final values of σ_X^2, σ_P^2 for a coherent state with a varied phase and a given $\bar{n}$ are found as averages over 10,000 points (that is, 40 runs like in Fig. 2b). For example, for $\bar{n} = 5$ we find $\sigma_{X(P)}^2 = 1.20(1.12)$ taken at gains 0.96 and 0.95 respectively. The results of $\sigma_{X,P}^2(g_{X,P})$ for a range of photon numbers $\bar{n} = 0$ (vacuum), $\bar{n} = 5, 20, 45, 180, 500$ at various

gains are summarized in a figure in the Supplementary Methods. From this we obtain $\sigma_{X,P}^2(g_{X,P})$, which can be inserted into the fidelity expression. For a given width of the gaussian distribution of coherent states we find the values of g_X, g_P , and the corresponding $\sigma_{X,P}^2(g_{X,P})$ which maximize the fidelity. We obtain the following fidelities for distributions with a width $\langle n \rangle = 2, 5, 10, 20, 200$: $F_2 = 0.64 \pm 0.02$; $F_5 = 0.60 \pm 0.02$; $F_{10} = 0.59 \pm 0.02$; $F_{20} = 0.58 \pm 0.02$; $F_{200} = 0.56 \pm 0.03$. The expression for the classical benchmark fidelity¹¹ $F_n^{\text{class}} = \frac{\langle n \rangle + 1}{2(\langle n \rangle + 1)}$ gives $F_2^{\text{class}} = 0.60$; $F_5^{\text{class}} = 0.545$; $F_{10}^{\text{class}} = 0.52$; $F_{20}^{\text{class}} = 0.51$; $F_{200}^{\text{class}} = 0.50$ (see Supplementary Methods for details on the fidelity calculations). The maximal $\langle n \rangle$ for successful teleportation is limited by small fluctuations of the classical gain, which for large $\bar{n}$ lead to large uncontrolled displacements of the teleported state with respect to the input state, and hence to the decrease in the fidelity.

In Fig. 3 we show the tomographically reconstructed teleported state with the mean photon number $\bar{n} = 5$. Owing to the gaussian character of the state, the knowledge of the means and the variances of two quadrature phase operators is sufficient for the reconstruction.

Note that the atomic object onto which the teleportation is performed contains hundreds of billions of atoms. However, the number of excitations in the ensemble, of course, corresponds to the number of photons in the initial state of light. Those excitations are coherently distributed over the entire ensemble.

Having demonstrated the teleportation for gaussian states, we now address the applicability of this teleportation protocol to the teleportation of a light qubit, which is relevant for, for example, quantum computing³. In the Supplementary Notes we give the derivation of the predicted qubit fidelity, F_q , based on the performance of our teleportation protocol for coherent states. For experimentally relevant values of losses and decoherence, $F_q = 0.72$ —higher than the best classical fidelity for a qubit of 0.67—can be predicted. In order to experimentally demonstrate such qubit teleportation, a source generating such a qubit in a temporal, spectral and spatial mode compatible with our atomic target is required. First steps towards generation of an atom-compatible qubit state of light have been recently made using atomic ensembles^{19–21}, single atoms in a cavity^{22,23}, and a photon subtracted squeezed state²⁴.

In our experiment, the entanglement generation and the Bell measurement overlap in time because the duration of the strong pulse and the pulse to be teleported is 2 ms, which is much longer than the time it takes light to travel from Alice to Bob. This situation, also the case in some teleportation experiments^{6,8}, is different, for example, from the teleportation^{7,9,10} in which the entanglement generation and the Bell measurement are separated in time. This feature is not inherent to our teleportation scheme—indeed, in principle, a shorter strong pulse (of higher power) would generate the same entanglement on a timescale short compared to the propagation time, especially if the distance from Alice to Bob is extended to a few kilometres. The teleportation distance can be increased, and is limited only by propagation losses of light and the atomic coherence lifetime. The timing of the entanglement generation and the Bell measurement may be potentially important for future applications.

Further improvement of the present teleportation protocol can be achieved by performing more complex photocurrent processing with the same homodyne set-up. As shown in ref. 12 and in the Supplementary Notes, a fidelity of 0.93 can be achieved if such processing is combined with the use of an experimentally feasible²⁵ 6 dB squeezed strong pulse.

METHODS

Calibration and measurement techniques. Physically, we perform measurements of the Stokes operators of light by two sets of balanced homodyne detectors (Fig. 1). The measurements on the first pulse represent the generalized Bell measurement. The same measurements on the second (verifying) pulse allow us to determine the teleported atomic state by performing quantum state

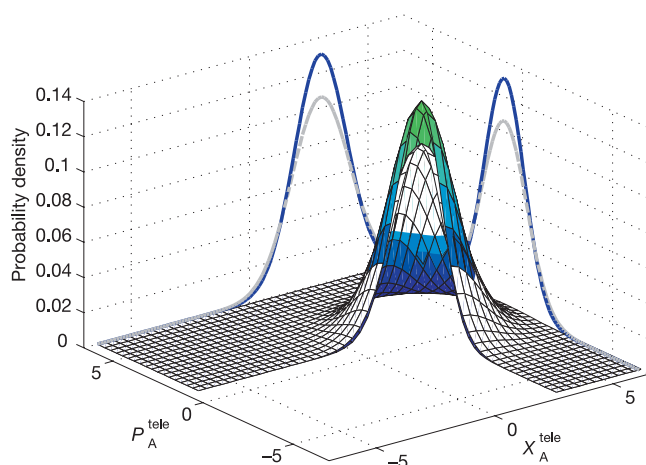


Figure 3 | Tomographic reconstruction of a teleported state with $\bar{n} = 5$ (coloured contour) versus the state corresponding to the best classical state transfer. Canonical variables plotted on horizontal axes are normalized so that their variance for a vacuum state is $1/2$.

tomography. The relevant $\cos(\Omega t)$ and $\sin(\Omega t)$ modulation components of the Stokes operators are measured by processing the corresponding photocurrents with lock-in amplifiers. The Stokes operators of interest are $\hat{S}_2$ (which is the difference between photon fluxes in the modes polarized at $\pm 45^\circ$ to the vertical axis, and $\hat{S}_3$ (which is the corresponding quantity for the left- and right-hand circular polarizations).

Calibration of the measurement of canonical variables for light is based on measurements of the shot (vacuum) noise level. We measure the Stokes parameters for the x -polarization mode in a vacuum state. The linear dependence of the variance of the measured photocurrents on the optical power of the strong pulse proves that the polarization state of light is, in fact, shot (vacuum) noise limited²⁵. All other measurements of $\hat{S}_2$, $\hat{S}_3$ are then normalized to this shot noise level, yielding the canonical variables as:

$$y_c = \frac{1}{\sqrt{2} \int_0^T d\tau \cos(\Omega t) S_2^{\text{vacuum}}(\tau)} \int_0^T d\tau \cos(\Omega t) S_2(\tau)$$

and similarly for $q_c(S_3)$ and the $\sin(\Omega t)$ components. Since our detectors have nearly unity (better than 0.97) quantum efficiency, the Stokes operators can be operationally substituted with measured photocurrents.

Next we need to calibrate the atomic coherent (projection) noise level. Whereas balanced homodyne detection for light has become an established technique for determination of the vacuum state⁶, a comparable technique for atoms is a relatively recent invention. Here we utilize the same procedure as used in our previous experiments on the atoms–light quantum interface^{14,15}. We use the fact that the vacuum (projection) noise level for collective atomic spin states in the presence of a bias magnetic field can be determined by sending a pulse of light through two identical atomic ensembles with oppositely oriented macroscopic spin orientation. We therefore insert a second atomic cell in the beam. As described in detail in ref. 15, the transmitted light state in this experiment is given by:

$$\hat{y}_c^{\text{out}} = \hat{y}_c^{\text{in}} + \frac{\kappa}{\sqrt{2}} (\hat{P}_{\text{atom1}} + \hat{P}_{\text{atom2}}) = \hat{y}_c^{\text{in}} + \kappa \hat{P}_{\text{total}}$$

where $\hat{P}_{\text{total}}$ is the spin canonical variable for the entire 2-cell atomic sample. Intuitively this equation can be understood by noting that terms proportional to κ^2 in equation (1) cancel out for propagation through two oppositely oriented ensembles. A similar equation holds for $\hat{y}_s^{\text{out}}$ with substitution of $\hat{X}_{\text{total}}$ for $\hat{P}_{\text{total}}$. The results for $\text{Var}(\hat{y}_{cs}^{\text{out}})$ as a function of the number of atoms are shown in the figure in the Supplementary Methods. The fact that the points lie on a straight line, along with the independent measurement of the degree of spin polarization above 0.99, proves^{14,15,18} that we are indeed measuring the vacuum (projection) noise of the atomic ensemble. κ^2 for different atomic numbers is then calculated from the graph (Supplementary Methods). Its values are in good agreement with the theoretical calculation¹⁸ according to $\kappa = a_1 \sqrt{N_{\text{ph}} N_{\text{at}} F \sigma \Gamma / A \Delta}$. In the experiment, we monitor the number of atoms by sending a weak off-resonant probe pulse along the direction x and measuring the Faraday rotation angle proportional to the collective macroscopic spin of the ensemble $J_x = 4N_{\text{at}}$. This Faraday angle is monitored throughout the teleportation experiment, so that the value of κ^2 is known at every stage.

Decoherence and losses. The main sources of imperfections are decoherence of the atomic state and reflection losses of light. For experimental values of the atomic decoherence and losses, the model developed in ref. 12 predicts, for example, $F_5 = 0.66$, which is still higher than the observed value owing to imperfections unaccounted for in the model but comparable to the obtained experimental results. Dissipation also affects the experimental state reconstruction procedure. The main effect of the light losses $\varepsilon = 0.09$ is that it modifies κ into $\kappa\sqrt{1-\varepsilon}$. However, this modified κ is, in fact, exactly the parameter measured in the two-cell calibration experiment described above, so no extra correction is due because of these losses. There is also a small amount of electronic noise of detectors which can be treated as an extra vacuum contribution to the input state.

Standard deviation of the teleportation fidelity. The standard deviation of the fidelity for $\langle n \rangle \leq 20$ is calculated as follows:

$$\begin{aligned} \text{SD}(F) &= \sqrt{\delta_{\text{PN}}^2 + \delta_{\text{SN}}^2 + \delta_{\text{el}}^2 + \delta_{\beta}^2 + \delta_{\text{SNR}}^2 + \delta_{\text{fit}}^2 + \delta_g^2} \\ &= 10^{-2} \sqrt{1.0^2 + 1.65^2 + 0.1^2 + 0.3^2 + 0.2^2 + 1.2^2 + 0.8^2} \approx 0.02 \end{aligned}$$

where $\delta_{\text{PN}} = 0.01$ is the contribution to the $\text{SD}(F)$ due to the projection noise fluctuations including an error due to imperfect optical pumping, $\delta_{\text{SN}} = 0.017$ is the contribution due to the shot noise level uncertainty, $\delta_{\text{el}} = 0.001$ is the contribution of the electronics noise level fluctuations, $\delta_{\beta} = 0.003$ is the uncertainty due to fluctuations in the atomic decay constant, $\delta_{\text{SNR}} = 0.002$ is the contribution of the fluctuations in the ratio of responses of two pairs of detectors, $\delta_{\text{fit}} = 0.012$ is the deviation due to the uncertainty of the quadratic fit

of the atomic noise as a function of gain, and $\delta_g = 0.008$ is the contribution of the gain fluctuations. For $\langle n \rangle > 20$, $\delta_{\text{fit}} = 0.016$, giving $\text{SD}(F) \approx 0.026 \approx 0.03$.

Received 6 May; accepted 28 July 2006.

- Bennett, C. H. *et al.* Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels. *Phys. Rev. Lett.* **70**, 1895–1899 (1993).
- Briegel, H. J., Dur, W., Cirac, J. I. & Zoller, P. Quantum repeaters: the role of imperfect local operations in quantum communication. *Phys. Rev. Lett.* **81**, 5932–5935 (1998).
- Gottesman, D. & Chuang, I. Demonstrating the viability of universal quantum computation using teleportation and single-qubit operations. *Nature* **402**, 390–393 (1999).
- Bouwmeester, D. *et al.* Experimental quantum teleportation. *Nature* **390**, 575–579 (1997).
- Boschi, D., Branca, S., De Martini, F., Hardy, L. & Popescu, S. Experimental realization of teleporting an unknown pure quantum state via dual classical and Einstein-Podolsky-Rosen channels. *Phys. Rev. Lett.* **80**, 1121–1125 (1998).
- Furusawa, A. *et al.* Unconditional quantum teleportation. *Sci. Tech. Froid* **282**, 706–709 (1998).
- de Riedmatten, H. *et al.* Long distance quantum teleportation in a quantum relay configuration. *Phys. Rev. Lett.* **92**, 047904 (2004).
- Takei, N., Yonezawa, H., Aoki, T. & Furusawa, A. High-fidelity teleportation beyond the no-cloning limit and entanglement swapping for continuous variables. *Phys. Rev. Lett.* **94**, 220502 (2005).
- Barrett, M. D. *et al.* Deterministic quantum teleportation of atomic qubits. *Nature* **429**, 737–739 (2004).
- Riebe, M. *et al.* Deterministic quantum teleportation with atoms. *Nature* **429**, 734–737 (2004).
- Hammerer, K., Wolf, M. M., Polzik, E. S. & Cirac, J. I. Quantum benchmark for storage and transmission of coherent states. *Phys. Rev. Lett.* **94**, 150503 (2005).
- Hammerer, K., Polzik, E. S. & Cirac, J. I. Teleportation and spin squeezing utilizing multimode entanglement of light with atoms. *Phys. Rev. A* **72**, 052313 (2005).
- Vaidman, L. Teleportation of quantum states. *Phys. Rev. A* **49**, 1473–1476 (1994).
- Julsgaard, B., Sherson, J., Fiurásek, J., Cirac, J. I. & Polzik, E. S. Experimental demonstration of quantum memory for light. *Nature* **432**, 482–486 (2004).
- Julsgaard, B., Kozhekin, A. & Polzik, E. S. Experimental long-lived entanglement of two macroscopic objects. *Nature* **413**, 400–403 (2001).
- Julsgaard, B., Schori, C., Sørensen, J. L. & Polzik, E. S. Atomic spins as a storage medium for quantum fluctuations of light. *Quant. Inf. Comput.* **3** (special issue), 518–534 (2003).
- Julsgaard, B., Sherson, J., Sørensen, J. L. & Polzik, E. S. Characterizing the spin state of an atomic ensemble using the magneto-optical resonance method. *J. Opt. B* **6**, 5–14 (2004).
- Sherson, J., Julsgaard, B. & Polzik, E. S. Deterministic atom-light quantum interface. *Adv. At. Mol. Opt. Phys.* (in press); preprint at (<http://arxiv.org/quant-ph/0601186>) (2006).
- Chou, C. W., Polyakov, S. V., Kuzmich, A. & Kimble, H. J. Single-photon generation from stored excitation in an atomic ensemble. *Phys. Rev. Lett.* **92**, 213601 (2004).
- Chaneliere, T. *et al.* Storage and retrieval of single photons transmitted between remote quantum memories. *Nature* **438**, 833–836 (2005).
- Eisaman, M. D. *et al.* Electromagnetically induced transparency with tunable single-photon pulses. *Nature* **438**, 837–841 (2005).
- Kuhn, A., Hennrich, M. & Rempe, G. Deterministic single-photon source for distributed quantum networking. *Phys. Rev. Lett.* **89**, 067901 (2002).
- McKeever, J. *et al.* Deterministic generation of single photons from one atom trapped in a cavity. *Science* **303**, 1992–1994 (2004).
- Neergaard-Nielsen, J. S., Melholt Nielsen, B., Hettich, C., Mølmer, K. & Polzik, E. S. Generation of a superposition of odd photon number states for quantum information networks. *Phys. Rev. Lett.* **97**, 083604 (2006).
- Polzik, E. S., Carri, J. & Kimble, H. J. Spectroscopy with squeezed light. *Phys. Rev. Lett.* **68**, 3020–3023 (1992).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements The experiment was performed at the Niels Bohr Institute, and was funded by the Danish National Research Foundation through the Center for Quantum Optics (QUANTOP), by EU grants COVAQIAL and QAP, and by the Carlsberg Foundation. I.C. and E.S.P. acknowledge the hospitality of the Institut de Ciències Fotòniques (ICFO) in Barcelona where ideas leading to this work were first discussed. The permanent address of K.H. is the Institut für theoretische Physik, Innsbruck, Austria.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to E.S.P. (polzik@nbi.dk).

Rapid subtropical North Atlantic salinity oscillations across Dansgaard–Oeschger cycles

Matthew W. Schmidt¹†, Maryline J. Vautravers²† & Howard J. Spero¹

Geochemical and sedimentological evidence suggest that the rapid climate warming oscillations of the last ice age, the Dansgaard–Oeschger cycles¹, were coupled to fluctuations in North Atlantic meridional overturning circulation through its regulation of poleward heat flux². The balance between cold meltwater from the north and warm, salty subtropical gyre waters from the south influenced the strength and location of North Atlantic overturning circulation during this period of highly variable climate^{3–5}. Here we investigate how rapid reorganizations of the ocean–atmosphere system across these cycles are linked to salinity changes in the subtropical North Atlantic gyre. We combine Mg/Ca palaeothermometry and oxygen isotope ratio measurements on planktonic foraminifera across four Dansgaard–Oeschger cycles (spanning 45.9–59.2 kyr ago) to generate a seawater salinity proxy record from a subtropical gyre deep-sea sediment core. We show that North Atlantic gyre surface salinities oscillated rapidly between saltier stadial conditions and fresher interstadials, covarying with inferred shifts in the Tropical Atlantic hydrologic cycle⁶ and North Atlantic overturning circulation. These salinity oscillations suggest a reduction in precipitation into the North Atlantic and/or reduced export of deep salty thermohaline waters during stadials. We hypothesize that increased stadial salinities preconditioned the North Atlantic Ocean for a rapid return to deep overturning circulation and high-latitude warming by contributing to increased North Atlantic surface-water density on interstadial transitions.

Through its influence on regional evaporation–precipitation patterns in the North Atlantic basin^{3,7,8}, atmospheric freshwater transport plays an important part in maintaining the increased North Atlantic gyre surface salinities that promote meridional overturning circulation (MOC). Net evaporation exceeds precipitation in the western subtropical North Atlantic gyre, increasing the surface salinity of the Gulf Stream as it circulates northward through the subtropics^{3,9}. After these surface waters reach subpolar latitudes and warm the atmosphere, the excess salt is critical in maintaining modern North Atlantic MOC (ref. 10). It is generally agreed that rapid climate shifts associated with Dansgaard–Oeschger cycles during the last glacial period were triggered by changes in surface water density at the sites of deep water formation in the North Atlantic^{2,4,5}. Whereas warm interstadials were characterized by a circulation pattern similar to that of the present, the northward flow of warm, salty surface waters into the subpolar North Atlantic was reduced and/or subducted below a cold, fresh surface layer during stadials^{4,5}. As a result, stadial surface temperatures plummeted in the northern North Atlantic region and the sites of deep-water formation migrated south of the stadial polar front (Fig. 1).

We reconstruct surface salinities (as seawater $\delta^{18}\text{O}$, $\delta^{18}\text{O}_{\text{SW}}$) across four Dansgaard–Oeschger cycles at Ocean Drilling Program (ODP)

site 1060 located on the Blake Outer Ridge near the modern position of the northward flowing Gulf Stream (Fig. 1). Seawater $\delta^{18}\text{O}$ covaries linearly with surface salinity, making it one of the most direct proxies for estimating salinity in the modern ocean¹¹. Because foraminiferal calcite ($\delta^{18}\text{O}_{\text{C}}$) is controlled by temperature and $\delta^{18}\text{O}_{\text{SW}}$, $\delta^{18}\text{O}_{\text{SW}}$ can be computed if temperature is determined independently. Using combined Mg/Ca palaeothermometry and $\delta^{18}\text{O}$ analyses of shells from the surface-dwelling foraminifera *Globigerinoides ruber* (white variety), we produce a continuous millennial-scale record (4 cm sampling, 75–250-year resolution) of gyre $\delta^{18}\text{O}_{\text{SW}}$ variability. We use previously published methods^{12,13} for measurements and for calculations of $\delta^{18}\text{O}_{\text{SW}}$. Our results are presented on a previously published age model developed for site 1060 based on correlations between the percentage of warm surface-dwelling planktonic foraminifera and changes in atmospheric temperature over Greenland as determined from the Greenland Icecore Project (GRIP) ice $\delta^{18}\text{O}$ record¹⁴ (see Supplementary Figs 1 and 2).

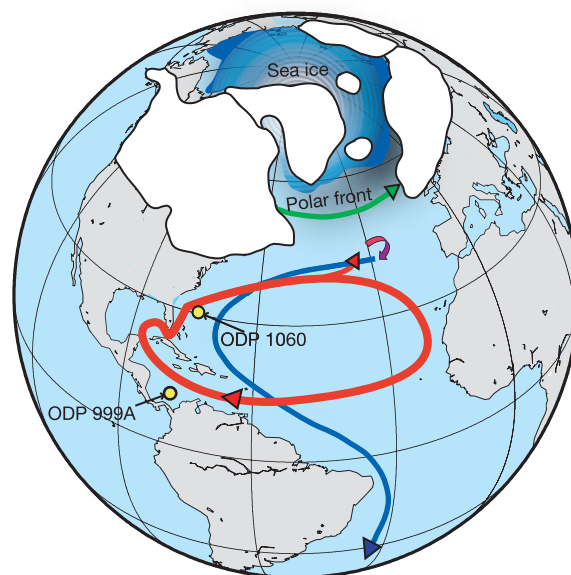


Figure 1 | Idealized North Atlantic surface (red arrows) and deep water (blue arrow) circulation during MIS3 stadial events. Flow paths are based on modern patterns. The locations of ODP site 1060 (30° 46' N, 74° 28' W; 3,480 m; 20–53 cm kyr^{−1} sedimentation rate) and ODP site 999A (12° 45' N, 78° 44' W) are indicated. The stadal mode of overturning circulation results from the convection of surface waters south of the polar front. The extent of continental ice sheets is estimated from ref. 30.

¹Department of Geology, University of California, Davis, California 95616, USA. ²Godwin Laboratory for Palaeoclimate Research, University of Cambridge, Cambridge CB2 3EQ, UK. †Present addresses: School of Earth and Atmospheric Sciences, Georgia Institute of Technology, Atlanta, Georgia 30332, USA (M.W.S.); British Antarctic Survey, High Cross, Madingley Road, Cambridge CB3 0ET, UK (M.J.V.).

The *G. ruber* $\delta^{18}\text{O}_\text{C}$ record from site 1060 displays clear isotopic oscillations with the lowest values occurring during warm interstadial events and the highest values occurring during cold stadials (Fig. 2a). Average stadial–interstadial amplitudes are $\sim 0.7\text{‰}$. Interstadial $\delta^{18}\text{O}_\text{C}$ values are lowest during the middle of interstadials and gradually increase during the latter part of warm phases.

Mg/Ca ratios in *G. ruber* were converted to sea surface temperature (SST) using a species-specific, depth-corrected relationship calibrated for the Atlantic¹⁵. Mg/Ca SSTs from site 1060 (Fig. 2b) do not mirror the GRIP ice $\delta^{18}\text{O}$ record (Fig. 2e). Whereas the transitions into interstadials 12, 13, 15, and 16 show a SST increase, temperature change at the transition into interstadial 14 is indistinct. Although Mg/Ca SSTs during the latter part of interstadials 14 and 16 are similar to or lower than faunal-based August SST calculations¹⁴, Mg/Ca SSTs are warmer than August faunal temperatures during the remaining interstadials and are as much as 5°C warmer than August faunal SSTs during stadials (Fig. 2b). Mg/Ca and Sr/Ca ratios in *G. ruber* shell material from site 1060 do not covary and Mg/Ca ratios do not correlate with a nearby benthic *Cibicides wuellerstorfi* $\delta^{13}\text{C}$ record also on the Blake Outer Ridge, suggesting that dissolution and stadial–interstadial changes in bottom-water masses at site 1060 have not affected our reconstructed SST record (see Supplementary Information Figs 3 and 4).

Because foraminiferal productivity is seasonally biased, *G. ruber* and other species will not record annual mean SST^{16,17}. The observation that Mg/Ca SST reconstructions are similar to August faunal SST estimates at site 1060 during interstadials (Fig. 2b) suggests that *G. ruber* inhabited the water column above site 1060 during the warmest summer months of marine isotope stage (MIS) 3. In addition, the largest temperature differences between the faunal and the Mg/Ca SST reconstructions occur during stadials. One possible explanation for the offset between proxies is that a steep stadial temperature gradient existed in this region of the subtropical gyre. As a result, *G. ruber* would have been excluded from the faunal population above site 1060 during most of the year when the frontal boundary was located south of site 1060. However, when northward movement of the meandering Gulf Stream shifted the frontal boundary near site 1060 during the warmest years and summer months, subtropical gyre waters containing *G. ruber* would have provided a source for the flux of this species to the sediment. We therefore interpret the *G. ruber* geochemical record to reflect late summer conditions in the subtropical gyre. Most important for salinity reconstructions is the fact that $\delta^{18}\text{O}_\text{C}$ and Mg/Ca SST are measured on material from the same population of *G. ruber*. Hence, these two geochemical proxies are linked to the same physical properties of the surface ocean, thereby allowing us to compute and compare $\delta^{18}\text{O}_\text{SW}$ values at site 1060 when *G. ruber* was present.

Removing temperature from the $\delta^{18}\text{O}_\text{C}$ record using the temperature– $\delta^{18}\text{O}$ relationship that has been field-calibrated for use with white *G. ruber* (ref. 18) yields the salinity proxy $\delta^{18}\text{O}_\text{SW}$ (Fig. 2c). Unlike Mg/Ca SST, $\delta^{18}\text{O}_\text{SW}$ values at site 1060 show clear salinity shifts that relate to stadial–interstadial phases in the GRIP $\delta^{18}\text{O}$ ice record¹ (Fig. 2e). However, bioturbation and potential age model errors do not allow us to determine definitively the timing of these $\delta^{18}\text{O}_\text{SW}$ transitions relative to changes in the GRIP ice core. Nevertheless, stadial events are characterized by the abrupt development of saltier, more positive $\delta^{18}\text{O}_\text{SW}$ values that average between 2.2 and 2.3‰. In contrast, interstadials 16, 14, and 12 have $\delta^{18}\text{O}_\text{SW}$ values that average between 1.7 and 1.9‰. The previously reconstructed MIS 3 $\delta^{18}\text{O}_\text{SW}$ values in the Caribbean core ODP 999A are consistent with the average stadial–interstadial site 1060 $\delta^{18}\text{O}_\text{SW}$ values from 45 to 60 kyr (ref. 13) (see Supplementary Fig. 5).

Although the magnitude of sea level change during MIS 3 has yet to be determined precisely, changes in sea level cannot account for the 0.3–0.6‰ $\delta^{18}\text{O}_\text{SW}$ differences between stadial–interstadial phases (see Supplementary Fig. 6). We therefore conclude that the majority of the $\delta^{18}\text{O}_\text{C}$ variation seen in *G. ruber* is due to salinity change, and

consider these $\delta^{18}\text{O}_\text{SW}$ shifts a robust reflection of Gulf Stream/gyre surface salinity variability during MIS 3. If we assume that the $\delta^{18}\text{O}_\text{SW}$ –surface salinity relationship for the southwestern North Atlantic gyre during MIS 3 was similar to the modern relationship¹⁹,

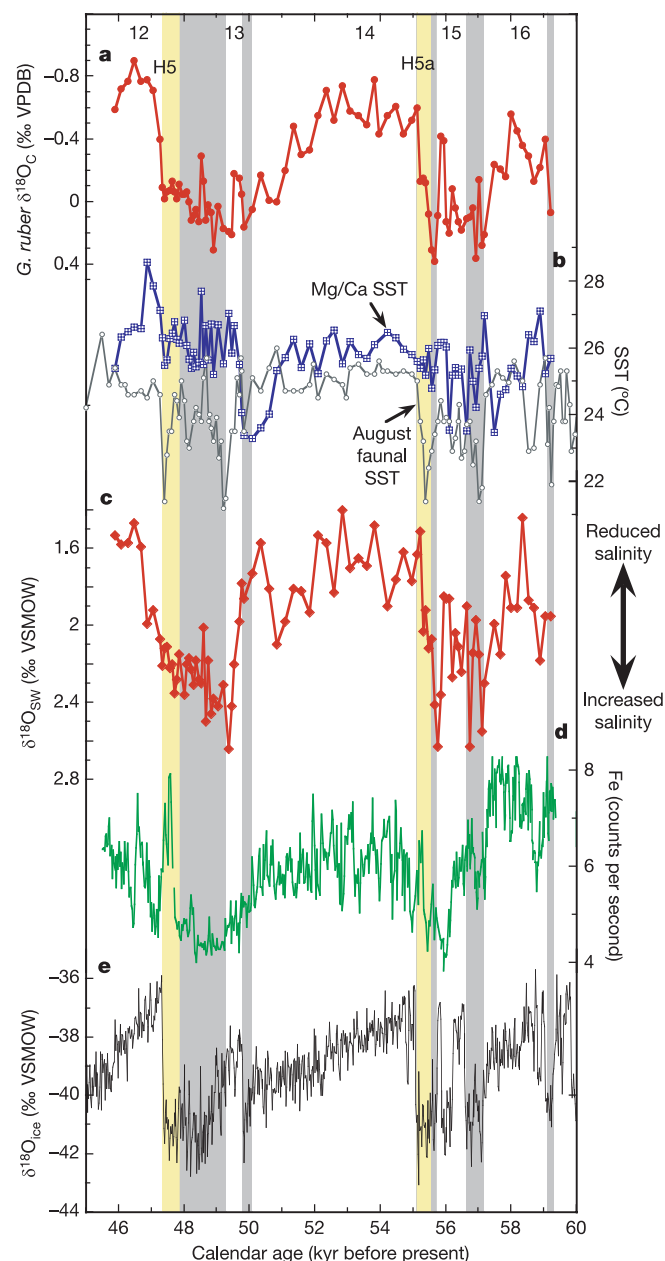


Figure 2 | Temperature and salinity variation at site 1060 during MIS 3. **a**, *G. ruber* (white variety) $\delta^{18}\text{O}_\text{C}$ from site 1060 (solid red circles). VPDB, Vienna Pee-Dee belemnite standard. **b**, Mg/Ca SSTs (blue crossed squares) from site 1060. Mg/Ca ratios were converted to SST using¹⁵ $\text{Mg}/\text{Ca} = 0.38 \times e^{0.09(\text{SST} - 0.61)}$, where SST is in $^\circ\text{C}$ and d is the depth of the core in kilometres. The faunal SST reconstructions for August (grey open circles) calculated from the foraminiferal assemblages¹⁴ are also shown for comparison. **c**, Computed $\delta^{18}\text{O}_\text{SW}$ from site 1060 (solid red diamonds) calculated from the Mg/Ca-derived SST and $\delta^{18}\text{O}_\text{C}$ using $T\text{ (in }^\circ\text{C)} = 16.5 - 4.80(\delta^{18}\text{O}_\text{C} - (\delta^{18}\text{O}_\text{SW} - 0.27\text{‰}))$. **d**, Fe content (counts per second) from the Cariaco basin core ODP site 1002 (ref. 6), adjusted to the GRIP age model. **e**, GRIP $\delta^{18}\text{O}_\text{ICE}$ record¹. Cold stadials in the GRIP record are marked by more arid conditions in the western tropical Atlantic (lower Fe concentrations) and elevated gyre salinities (increased $\delta^{18}\text{O}_\text{SW}$ values at site 1060). Note the salinity increases (as $\delta^{18}\text{O}_\text{SW}$) associated with the shaded stadial events and the reduced salinities associated with interstadials. Numbered interstadials 12 to 16 and Heinrich events H5 and H5a (yellow bars) are labelled at the top of the figure.

$\delta^{18}\text{O}_{\text{SW}} = 0.4 \times (\text{surface salinity}) - 13.5$, then average stadial–interstadial $\delta^{18}\text{O}_{\text{SW}}$ values suggest millennial-scale surface-salinity enrichments of 0.7 to 1.5 during stadials (see ‘Error analysis’ in Methods).

Changes in the Fe and Ti content in the Cariaco basin core ODP 1002 (ref. 6) provides a high-resolution record of high to low latitude air–sea linkages in the Atlantic during MIS 3. These data suggested that the intertropical convergence zone (ITCZ) migrated southward during cold stadial events in the North Atlantic, resulting in reduced freshwater precipitation over northeastern South America and reduced river-borne delivery of Fe and Ti to the Cariaco basin⁶. Although our site 1060 $\delta^{18}\text{O}_{\text{SW}}$ record is of lower resolution, wet/dry periods in the ODP site 1002 Fe record covary with changes in surface $\delta^{18}\text{O}_{\text{SW}}$ values at site 1060 (Fig. 2c and d). Wet interstadials 16, 14 and 12 correspond to periods of reduced $\delta^{18}\text{O}_{\text{SW}}$ values at site 1060. Likewise, the sustained period of arid conditions in the Cariaco basin after interstadial 14 (49.2–47.7 kyr) corresponds to enriched $\delta^{18}\text{O}_{\text{SW}}$ values in the gyre. Neither the Cariaco basin Fe record nor the site 1060 $\delta^{18}\text{O}_{\text{SW}}$ record show the predicted response to the interstadial 13 warming recorded in the GRIP ice core (Fig. 2).

During periods of high-latitude cooling, general circulation model (GCM) simulations predict increased northeast trade wind strength associated with an intensification of the subtropical high pressure located over the southwestern North Atlantic gyre^{20,21}. These conditions result in a precipitation deficit in the tropical and subtropical North Atlantic coupled with a freshening of the tropical and subtropical South Atlantic. Speleothem records from Brazil indicate that the South Atlantic experienced enhanced freshwater input during cool stadial events in the north²², just as our $\delta^{18}\text{O}_{\text{SW}}$ reconstruction from site 1060 indicates elevated salinities in the North Atlantic gyre. This dipole wet/dry relationship between the North and South Atlantic implies that the tropical hydrologic cycle helps to regulate the meridional salinity gradient in the Atlantic basin.

GCM modelling of North Atlantic MOC stability under glacial boundary conditions also suggests that the tropical Atlantic hydrologic cycle shifts to an enhanced evaporative regime, resulting in glacial North Atlantic MOC stabilization^{23,24}. The increased stadial gyre salinities that we record at site 1060 are therefore consonant with the stable mode of haline-driven glacial North Atlantic MOC found in these models^{23,24}. We hypothesize that the stadial mode of MOC, coupled with a southward shift of the ITCZ, resulted in reduced oceanic salt export and reduced tropical precipitation into the north Atlantic. Together, these led to the rapid accumulation of salt in the North Atlantic subtropical gyre on transitions into stadials. Then, as the low-salinity surface waters and sea ice retreated in the high-latitude North Atlantic on interstadial transitions⁵, the positive gyre salinity anomalies would have penetrated northward within the North Atlantic Drift, elevating surface water density in the Greenland–Iceland–Norwegian seas and the Labrador Sea and promoting the resumption of vigorous interstadial MOC.

Reconstruction of $\delta^{18}\text{O}_{\text{SW}}$ from the western equatorial Pacific also indicates elevated stadial surface salinities (ref. 25), suggesting a link between North Atlantic stadials and El Niño-like conditions in the tropical Pacific. Coupled GCM^{26,27} and reanalysis results²⁸ show that El Niño conditions also result in enhanced water vapour removal from the tropical Atlantic. In addition, historical data indicate that the mean position of the ITCZ is displaced southward during an El Niño phase, resulting in a rainfall deficit in Central America²⁹. If the freshwater perturbations associated with El Niño conditions persisted for longer periods (such as during MIS 3 stadials), this could reinforce the drying effects of a more southerly ITCZ and contribute to the elevated stadial salinities we record in the North Atlantic gyre at site 1060.

Stadial–interstadial reorganizations of the tropical Atlantic hydrologic cycle therefore had a significant impact on the freshwater balance of the North Atlantic subtropical gyre. As the ITCZ migrated

southward during stadials, enhanced water vapour removal from the North Atlantic led to the development of positive salinity anomalies in the subtropical gyre, contradicting the hypothesis that interstadials are periods of increased freshwater export from the Atlantic⁶. A shift to enhanced El Niño-like stadial circulation in the eastern tropical Pacific may also have contributed to increased tropical Atlantic aridity. If, as we predict, increased surface salinities in the western subtropical gyre are indicative of an overall enrichment in the North Atlantic gyre’s mixed-layer salinity during stadials, then the tropical hydrologic cycle acts as a negative feedback during weak phases of MOC, increasing North Atlantic surface water density and preconditioning the system for a return to an interstadial mode of strong MOC.

METHODS

Sample preparation. Sediment samples were collected from site 1060 at 4 cm resolution. Each core interval was disaggregated in ultrapure water, sieved and dried at room temperature. If possible, specimens of *G. ruber* (white variety) were then collected from the 250–350- μm size fraction and 25 shells were combined for each stable isotope analysis for downcore measurements to minimize intraspecific variation in shell geochemistry. These samples were sonicated in methanol for 5–10 s, roasted in vacuum at 375 °C for 30 min, and analysed on a Fisons Optima infrared mass spectrometer using an Isocarb common acid bath autocarbonate system at 90 °C at the University of California, Davis. During stadial intervals, *G. ruber* abundance was much lower, so it was necessary to collect all specimens >63 μm in size in certain intervals. In these intervals, all available specimens were pooled and gently crushed. A separate split of the homogenized sample (approximately 40 μg) was roasted and analysed for stable isotopes and the remainder of the shell material was analysed for trace and minor elements.

Elemental ratio measurement. Mg/Ca ratios were measured on foraminifera collected from the same population and size fraction that was used for the stable isotope analyses. Where possible, approximately 600 μg of material per sample (~55 shells) was cleaned for trace- and minor-element analysis without a DTPA step¹². Briefly, samples underwent a multi-step process consisting of initial rinses in ultrapure water, followed by treatments with hot reducing and oxidizing solutions, transfers into new acid-leached micro-centrifuge vials, and finally leaches with a dilute ultrapure acid solution. All sample cleaning was conducted in laminar flow benches under trace-metal-clean conditions. Samples were then dissolved and analysed on a Finnigan Element-2 inductively coupled plasma mass spectrometer at the University of California, Santa Barbara, using established procedures¹².

Elemental ratios of Fe/Ca and Al/Ca were used to monitor cleaning efficacy. Fe/Ca ratios remain <130 $\mu\text{mol mol}^{-1}$ and Al/Ca ratios remain <50 $\mu\text{mol mol}^{-1}$ throughout the record, indicating that Mg contamination associated with detrital sediment in cleaned samples was not a problem¹⁷. Although Mg/Ca does show a statistically significant correlation with Fe/Ca, levels of Fe are three orders of magnitude smaller than Mg in the cleaned samples, so Fe-bearing minerals not removed during the cleaning process cannot contribute significantly to the primary Mg/Ca signal in the foraminiferal calcite.

Error analysis. Analytical precision for the $\delta^{18}\text{O}_{\text{C}}$ measurements is better than $\pm 0.06\text{‰}$. The pooled standard deviation of replicate Mg/Ca analyses from site 1060 was $\pm 2.9\%$ (1 standard deviation (s.d.); degrees of freedom = 49; 51% of the intervals run in replicate). The s.d. for the $\delta^{18}\text{O}_{\text{SW}}$ residual was calculated to be $\pm 0.2\text{‰}$ (1 s.d.) using a Monte Carlo methodology that assumed a normal distribution in the $\delta^{18}\text{O}_{\text{C}}$ and Mg/Ca errors and in the Mg/Ca–SST and $\delta^{18}\text{O}_{\text{C}}$ –SST calibration precision.

The conversion of the $\delta^{18}\text{O}_{\text{SW}}$ proxy to salinity is difficult to constrain owing to uncertainties in the slope of the relationship between these two variables at times in the past. However, it is likely that the slope of this relationship was steeper during glacial times owing to the input of isotopically depleted fresh water, reducing the millennial-scale surface salinities change in the gyre to <1 salinity unit.

Received 19 December 2005; accepted 27 July 2006.

1. Johnsen, S. J. *et al.* Oxygen isotope and palaeotemperature records from six Greenland ice-core stations: Camp Century, Dye-3, GRIP, GISP2, Renland and NorthGRIP. *J. Quat. Sci.* **16**, 299–307 (2001).
2. Boyle, E. A. Is ocean thermohaline circulation linked to abrupt stadial/interstadial transitions? *Quat. Sci. Rev.* **19**, 255–272 (2000).
3. Broecker, W. S., Bond, G., Klas, M., Bonani, G. & Wolff, W. A salt oscillator in the glacial Atlantic? The concept. *Paleoceanography* **5**, 469–478 (1990).

4. Oppo, D. W. & Lehman, S. J. Suborbital timescale variability of North Atlantic deep water during the past 200,000 years. *Paleoceanography* **10**, 900–910 (1995).
5. Rasmussen, T. & Thomsen, E. The role of the North Atlantic Drift in the millennial timescale glacial climage fluctuations. *Palaeogeogr. Palaeoclim. Palaeoceanogr.* **210**, 101–116 (2004).
6. Peterson, L. C., Haug, G. H., Hughen, K. A. & Rohl, U. Rapid changes in the hydrologic cycle of the Tropical Atlantic during the last Glacial. *Science* **290**, 1947–1951 (2000).
7. Broecker, W. S. The salinity contrast between the Atlantic and Pacific during glacial time. *Paleoceanography* **4**, 207–212 (1989).
8. Zaucker, F. & Broecker, W. S. The influence of atmospheric moisture transport on the fresh water balance of the Atlantic drainage basin: general circulation model simulations and observations. *J. Geophys. Res.* **97**, 2765–2773 (1992).
9. Curry, R., Dickson, B. & Yashayaev, I. A change in the freshwater balance of the Atlantic Ocean over the past four decades. *Nature* **426**, 826–829 (2003).
10. Broecker, W. S. The great ocean conveyor. *Oceanography* **4**, 79–89 (1991).
11. Craig, H. & Gordon, L. I. in *Stable Isotopes in Oceanographic Studies and Paleotemperatures* (ed. Tongiorgi, E.) 9–130 (CNR, Pisa, 1965).
12. Lea, D. W., Pak, D. K. & Spero, H. J. Climate impact of Late Quaternary equatorial Pacific sea surface temperature variations. *Science* **289**, 1719–1724 (2000).
13. Schmidt, M. W., Spero, H. J. & Lea, D. W. Links between salinity variation in the Caribbean and North Atlantic thermohaline circulation. *Nature* **428**, 160–163 (2004).
14. Vautravers, M. J., Shackleton, N. J., Lopez-Martinez, C. & Grimalt, J. O. Gulf Stream variability during marine isotope stage 3. *Paleoceanography* **19**, PA2011, doi:10.1029/2003PA000966 (2004).
15. Dekens, P. S., Lea, D. W., Pak, D. K. & Spero, H. J. Core top calibration of Mg/Ca in tropical foraminifera: refining palaeotemperature estimation. *Geochem. Geophys. Geosyst.* **4**, doi:10.1029/2001GC000200 (2002).
16. Deuser, W. G., Ross, E. H., Hemleben, C. & Spindler, M. Seasonal changes in species composition, numbers, mass, size, and isotopic composition of planktonic-foraminifera settling into the deep Sargasso Sea. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **33**, 103–127 (1981).
17. Schmidt, M. W., Vautravers, M. J. & Spero, H. J. Western Caribbean sea surface temperatures during the Late Quaternary. *Geochem. Geophys. Geosyst.* **7**, doi:10.1029/2005GC000957 (2006).
18. Thunell, R., Tappa, E., Pride, C. & Kincaid, E. Sea-surface temperature anomalies associated with the 1997–1998 El Niño recorded in the oxygen isotope composition of planktonic foraminifera. *Geology* **27**, 843–846 (1999).
19. Schmidt, G. A., Bigg, G. R. & Rohling, E. J. *Global Seawater Oxygen-18 Database* (<http://data.giss.nasa.gov/o18data/>) (1999).
20. Lohmann, G. Atmospheric and oceanic freshwater transport during weak Atlantic overturning circulation. *Tellus A* **55**, 438–449 (2003).
21. Vellinga, M. & Wu, P. L. Low-latitude freshwater influence on centennial variability of the Atlantic thermohaline circulation. *J. Clim.* **17**, 4498–4511 (2004).
22. Wang, X. F. et al. Wet periods in northeastern Brazil over the past 210 kyr linked to distant climate anomalies. *Nature* **432**, 740–743 (2004).
23. Prange, M., Romanova, V. & Lohmann, G. The glacial thermohaline circulation: stable or unstable? *Geophys. Res. Lett.* **29**, doi:10.1029/2002GL015337 (2002).
24. Romanova, V., Prange, M. & Lohmann, G. Stability of the glacial thermohaline circulation and its dependence on the background hydrological cycle. *Clim. Dyn.* **22**, 527–538 (2004).
25. Stott, L. D., Poulsen, C., Lund, S. & Thunell, R. Super ENSO and global climate oscillations at millennial time scales. *Science* **297**, 222–226 (2002).
26. Latif, M., Roeckner, E., Mikolajewicz, U. & Voss, R. Tropical stabilization of the thermohaline circulation in a greenhouse warming simulation. *J. Clim.* **13**, 1809–1813 (2000).
27. Schmittner, A. & Clement, A. C. Sensitivity of the thermohaline circulation to tropical and high latitude freshwater forcing during the last glacial–interglacial cycle. *Paleoceanography* **17**, doi:10.1029/2000PA (2002).
28. Schmittner, A., Appenzeller, C. & Stocker, T. F. Enhanced Atlantic freshwater export during El Niño. *Geophys. Res. Lett.* **27**, 1163–1166 (2000).
29. Hastenrath, S. The intertropical convergence zone of the eastern Pacific revisited. *Int. J. Climatol.* **22**, 347–356 (2002).
30. Peltier, W. R. Ice Age paleotopography. *Science* **265**, 195–201 (1994).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank the Ocean Drilling Program for core samples. Funding for this research was provided by JOI-ODP Schlanger Ocean Drilling Fellowships to M.W.S. and by a US National Science Foundation grant to H.J.S. Laboratory assistance from D. Pak and mass spectrometer operation by G. Paradis and D. Winter were critical to the success of this study. Elemental analyses were conducted in D. Lea's laboratory at the University of California, Santa Barbara. We also thank A. Russell and D. Lea for discussions and comments.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to M.W.S. (mschmidt@eas.gatech.edu).

Lithium isotope evidence for subduction-enriched mantle in the source of mid-ocean-ridge basalts

Tim Elliott¹, Alex Thomas^{1,2}, Alistair Jeffcoate^{1†} & Yaoling Niu³

'Recycled' crustal materials, returned from the Earth's surface to the mantle by subduction, have long been invoked to explain compositional heterogeneity in the upper mantle¹. Yet increasingly, problems have been noted with this model^{2,3}. The debate can be definitively addressed using stable isotope ratios, which should only significantly vary in primitive, mantle-derived materials as a consequence of recycling. Here we present data showing a notable range in lithium isotope ratios in basalts from the East Pacific Rise, which correlate with traditional indices of mantle heterogeneity (for example, ¹⁴³Nd/¹⁴⁴Nd ratios). Such co-variations of stable and radiogenic isotopes in melts from a normal ridge segment provide critical evidence for the importance of recycled material in generating chemical heterogeneity in the upper mantle. Contrary to many models, however, the elevated lithium isotope ratios of the 'enriched' East Pacific Rise lavas imply that subducted ocean crust is not the agent of enrichment. Instead, we suggest that fluid-modified mantle, which is enriched during residency in a subduction zone, is mixed back into the upper mantle to cause compositional variability.

The Li isotope system can provide a novel perspective on the role of recycled material in the mantle, but is yet to be fully exploited in studying geodynamic problems⁴. Subduction delivers altered oceanic crust and underlying serpentinized peridotite with heavy Li (that is, elevated ⁷Li/⁶Li or high $\delta^7\text{Li}$) to the mantle^{5,6}. Heavy $\delta^7\text{Li}$ in mantle-derived melts should therefore be a diagnostic tracer of recycled material. This simple scenario is made more interesting by processes that occur during subduction: altered oceanic crust loses its isotopically heavy Li to the overlying mantle in the subduction zone^{7,8} and ultimately becomes isotopically light as a consequence of its dehydration^{9,10}. Addition of recycled material can hence lower or raise the Li isotope ratio of the mantle, depending on whether the oceanic crust itself or subduction-modified mantle is the main agent of 'enrichment' (Fig. 1). Here we present the first Li isotope data on mid-ocean-ridge basalts (MORB) with suitable precision to identify possible recycled contributions.

We have analysed ten well-characterized samples^{11,12} from two locations ($\sim 10^\circ 30' \text{ N}$ and $\sim 11^\circ 20' \text{ N}$) on a morphologically typical ridge segment of the East Pacific Rise (EPR). The location at $\sim 11^\circ 20' \text{ N}$ is marked by an anomalously large number of 'enriched' MORB¹¹. Similarly enriched samples from $11^\circ 30' \text{ N}$ and $12^\circ 50' \text{ N}$ EPR were argued to be derived from mantle-containing recycled oceanic crust¹³. Samples from $\sim 10^\circ 30' \text{ N}$ are more normal, 'depleted' MORB. All samples provided sufficient material to allow rigorous hand-picking of pristine glass. Li isotopes were measured on 2–10 mg samples of glass, processed by two-stage separation chemistry and analysed by multi-collector, inductively coupled plasma mass spectrometry¹⁴. Each separated sample was analysed at least twice and showed a reproducibility that was typically better than

$\pm 0.15\text{‰}$ (2 standard deviations, 2 s.d.; Table 1 and Supplementary Information). Full repeat measurements of two samples for which there was sufficient glass and triplicate analyses of international standard BHVO-2 (Table 1 and Supplementary Information) reproduced results that were in-line with our long-term estimate of repeatability, $\pm 0.3\text{‰}$ (2 s.d.)¹⁴. Our analytical procedure allowed small variations in Li isotope ratios to be detected: these variations

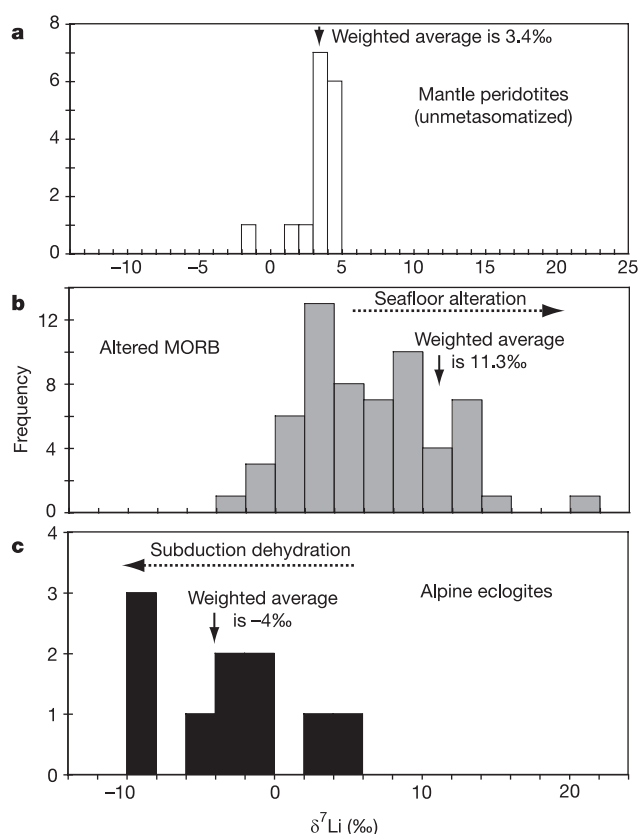


Figure 1 | Li isotopic compositions of the mantle and important recycled components. Histograms of $\delta^7\text{Li}$ for: **a**, modally unmetasomatized peridotites^{17–19}, taken to represent the most pristine shallow mantle (peridotite compositions are reconstituted from modal abundances and analyses of their constituent minerals for refs 17 and 18, and from bulk analyses for ref. 19); **b**, MORB altered by low-temperature interaction with sea water^{5,20}; and **c**, Alpine eclogites^{9,19}—tectonically exhumed fragments of mafic material interpreted to represent dehydrated, deep subducted oceanic crust. The average values reported in the figures were calculated by weighting individual isotopic analyses according to their Li concentrations.

¹Bristol Isotope Group, Department of Earth Sciences, Wills Memorial Building, Queen's Road, University of Bristol, Bristol BS8 1RJ, UK. ²Department of Earth Sciences, South Parks Road, Oxford University, Oxford OX1 3PR, UK. ³Department of Earth Sciences, Durham University, Durham DH1 3LE, UK. †Present address: Rio Tinto Iron Ore Atlantic, PO Box 695, 8th Floor, Castlemead, Lower Castle Street, Bristol BS99 1F5, UK.

would have been unresolved in previous Li isotope studies of MORB^{5,15,16}, which had uncertainties of ± 0.7 – 1.5‰ (2 s.d.; Fig. 2a). We report our Li isotope analyses in Table 1 as $\delta^7\text{Li}$, that is, the parts per thousand variation of $^7\text{Li}/^6\text{Li}$ relative to the standard NIST SRM 8545 (L-SVEC).

The $\delta^7\text{Li}$ of our samples ranges from 3.1 to 5.2‰ (Fig. 2) and so we are able to document Li isotope heterogeneity well beyond analytical uncertainty. The normal, depleted MORB samples have the lowest $\delta^7\text{Li}$, with values similar to the average of unmetasomatized mantle peridotites^{17–19} (Fig. 1a). Notably, our MORB Li isotope ratios show good correlations with a range of conventional elemental and isotopic indices of enrichment (Fig. 2). Heavier Li isotope ratios are associated with higher incompatible element concentrations (Fig. 2a), elevated ratios of highly incompatible to moderately

incompatible elements (Fig. 2b), distinctive highly incompatible element ratios (for example, Fig. 2c), more radiogenic $^{87}\text{Sr}/^{86}\text{Sr}$ (Fig. 2d) and less radiogenic $^{143}\text{Nd}/^{144}\text{Nd}$ (Fig. 2e).

Variations in the stable isotope ratios of mantle-derived rocks are an unequivocal signature of recycled material in their sources, provided they have not been modified by processes *en route* to the surface. It is thus necessary to assess processes that may have perturbed the primary $\delta^7\text{Li}$ of the magmas. Low-temperature alteration of the upper oceanic crust generates secondary minerals with high Li concentrations and heavy $\delta^7\text{Li}$ (refs 5, 20). Assimilation of this material could increase Li isotope ratios but should little affect absolute or relative abundances of elements unsusceptible to alteration and so cannot account for the trends observed in Figs 2b, c, e. Moreover, there is no correlation of $1/\text{Li}$ with $\delta^7\text{Li}$ in our data set (not shown) and the full range of our observed Li isotope variation is defined by two samples (PH 108-1 and PH 54-3) that are comparably differentiated (for example, 6.5 versus 7.0 wt% MgO) and have similar Li concentrations (6.8 and $6.9 \mu\text{g g}^{-1}$). Thus, heavy Li isotope ratios cannot represent the effects of preferential assimilation of Li-rich material at shallow levels. Mixing of more primitive magmas with anatectic melts of the magma chamber roof²¹ can potentially generate secondary variations in both abundances and ratios of 'fluid-immobile' elements and even $^{87}\text{Sr}/^{86}\text{Sr}$, but cannot account for changes in $^{143}\text{Nd}/^{144}\text{Nd}$. Therefore, interaction of mantle-derived melts with the crust seems unable to account for the geochemical variations we observe in the northern EPR.

It is also important to consider possible high-temperature fractionation of Li. Equilibrium isotopic fractionation of Li between melt and typical mantle peridotite is minor ($<0.5\text{‰}$)¹⁸. Likewise, differentiation of basaltic melts in Hawaii has been shown to cause no

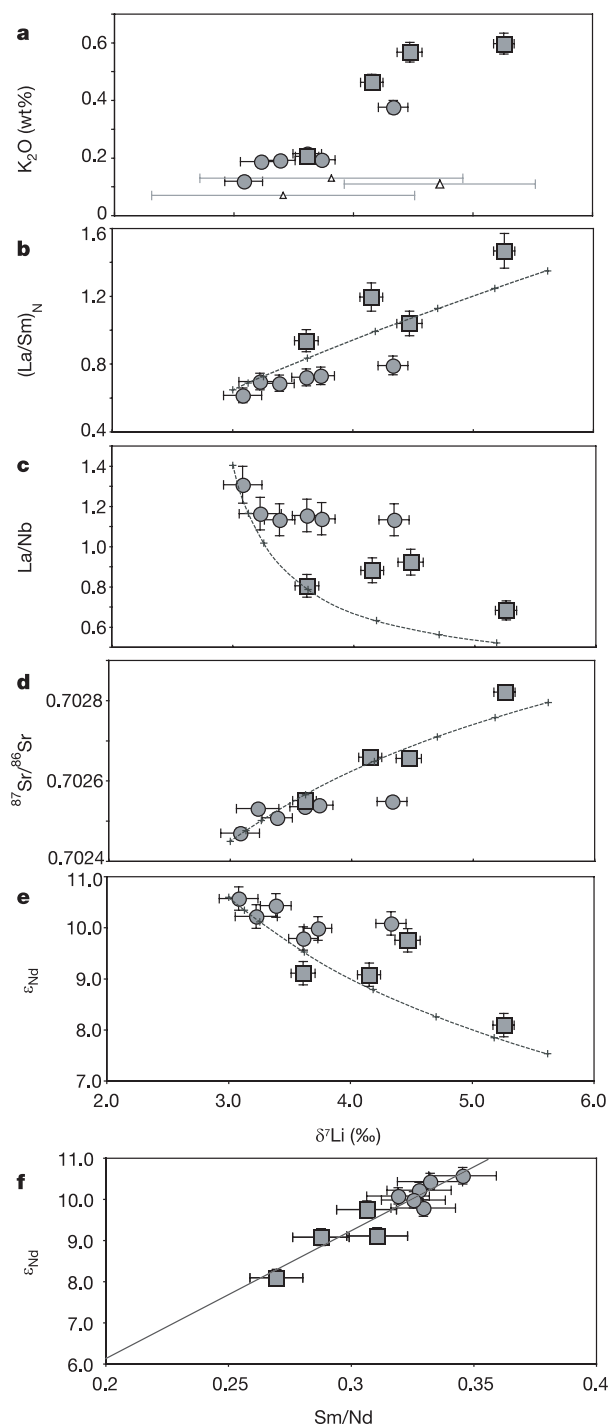


Figure 2 | Variations of $\delta^7\text{Li}$ with other geochemical parameters in northern EPR MORB. **a–e**, $\delta^7\text{Li}$ is plotted against: **a**, K_2O (wt%), together with previously published MORB⁵ analysed for both Li isotopes and K_2O (open triangles); **b**, chondrite normalized $^{143}\text{Nd}/^{144}\text{Nd}$; **c**, La/Nb (weight ratio); **d**, $^{87}\text{Sr}/^{86}\text{Sr}$; and **e**, ϵ_{Nd} (parts per 10^4 relative deviation of $^{143}\text{Nd}/^{144}\text{Nd}$ from the present chondrite reference value of 0.512638). **f**, Sm/Nd (weight ratio) is plotted against ϵ_{Nd} , together with a best fit line, which can be used to calculate an apparent age of enrichment (400 Myr ago). The $10^\circ 30'$ N samples are plotted as filled circles and $11^\circ 20'$ N samples are shown as filled squares. Apart from Li isotope measurements, the data are from refs 11, 12. The dashed lines in **b–e** are mixing curves between a hypothetical depleted upper mantle source with $\delta^7\text{Li} = 3\text{‰}$, $[\text{Li}] = 1.5 \mu\text{g g}^{-1}$, $[\text{Nb}] = 0.29 \mu\text{g g}^{-1}$, $[\text{La}] = 0.4 \mu\text{g g}^{-1}$, $[\text{Nd}] = 1.4 \mu\text{g g}^{-1}$, $[\text{Sm}] = 0.55 \mu\text{g g}^{-1}$, $[\text{Sr}] = 14 \mu\text{g g}^{-1}$ and the amphibole peridotite BZ72 (ref. 8), proposed to represent subduction-modified mantle. The composition of the depleted mantle is modelled by assuming that the most depleted sample of this study (PH 54-3) is the product of 10% mantle melting. Crosses indicate 0, 1, 2, 5, 10, 15, 20 and 25% (by weight) additions of BZ72 to the depleted mantle composition. Because there are no radiogenic isotope data reported for BZ72, we assume reasonable values of $^{87}\text{Sr}/^{86}\text{Sr} = 0.703$ and $\epsilon_{\text{Nd}} = 5.9$. The $^{87}\text{Sr}/^{86}\text{Sr}$ comes from an analysis of clinopyroxene separates from a similar, pargasite-bearing, Zarbagad peridotite—W36 (ref. 30). Given the low Rb/Sr of BZ72 (0.017) it is appropriate to use the $^{87}\text{Sr}/^{86}\text{Sr}$ value of a low Rb/Sr phase such as clinopyroxene. We inferred the Nd isotope value of the enriched endmember using the well-defined relationship in **f** and the measured Sm/Nd of BZ72 (0.194). Hence, there is some circularity in the fit of the mixing line in **e**, but since the other parameters used in its calculation are independent, the shape of the array still has some significance. The mixing trends shown in **b–e** serve to illustrate the general effects of addition of a single example of subduction-modified mantle to a normal MORB source, but were not necessarily expected closely to mimic the trends here. We did not model K_2O contents, which depend much more strongly than isotope and incompatible element ratios on assumptions about melting and fractionation, but note that BZ72 is a K_2O -rich peridotite (0.11 wt%). Error bars show 2 s.d. of repeat sample measurements (a conservative reproducibility of 6% for K_2O abundances was used, otherwise errors are documented in Table 1 and ref. 11).

Table 1 | $\delta^7\text{Li}$ measurements on EPR glasses

Sample	Latitude (°N)	Longitude (°W)	Depth (m)	MgO (wt%)	K ₂ O (wt%)	Li ($\mu\text{g g}^{-1}$)	$\delta^7\text{Li}$ (‰)	Error 2 s.d.	n
PH78-2	10.43	103.86	3058	3.9	0.38	13.1	4.3	0.1	4
PH77-6	10.43	103.85	3051	5.8	0.19	9.4	3.4	0.1	2
PH65-1	10.45	103.64	2996	5.5	0.19	10.6	3.2	0.2	4
PH64-2	10.46	103.63	3008	5.5	0.22	8.6	3.6	0.1	3
PH62-1	10.48	103.63	2812	6.1	0.19	8.2	3.7	0.1	4
PH54-3	10.50	103.51	3069	7.0	0.12	6.9	3.1	0.2	4
PH108-1	11.34	103.79	2724	6.5	0.60	6.8	5.2	0.1	4
PHGC-60	11.35	103.77	2545	6.0	0.46	7.9	4.2	0.1	4
Repeat							3.9	0.1	2
PH103-2	11.37	103.78	2570	4.8	0.57	9.9	4.5	0.1	3
PH94-1	11.37	103.71	2863	7.4	0.21	6.3	3.6	0.1	4
Repeat							3.4	0.1	2
BHVO-2							4.8	0.3	4
Repeat							4.7	0.1	1
Repeat							4.7	0.1	3

Errors reported as 2 s.d. of repeat measurements of the sample solution (n is number of repeat measurements). All individual measurements are documented in Supplementary Information. One BHVO-2 replicate was run only once and so the standard error of the run rather than the standard deviation of repeats is reported (in italics). Analyses of full repeat dissolutions (Repeat) of samples are reported on separate lines. Additional information on the samples was obtained from refs 11, 12. $\delta^7\text{Li} = [(^7\text{Li}/^6\text{Li})_{\text{sample}} / (^7\text{Li}/^6\text{Li})_{\text{NIST SRM8545}} - 1] \times 1000$.

systematic change in $\delta^7\text{Li}^{22}$. Finally, it has been argued that a diffusive mechanism of high-temperature Li isotopic fractionation could increase the $\delta^7\text{Li}$ of erupted magmas as a result of melt–peridotite interaction²³. Whether such a process typically operates beneath mid-ocean ridges remains unclear. Moreover, any variations of $\delta^7\text{Li}$ thus produced are anticipated to be correlated with Li abundances, which is not the case.

Hence, we are confident that the heavy Li isotope signatures of the enriched MORB are a consequence of recycled material in their sources. This is a highly significant finding in itself but the heavy Li isotope signatures of enriched MORB provide additional constraints on the origin of the recycled component. Because dehydrated residues of deep, subducted oceanic crust are isotopically light^{9,19}, recycled oceanic crust itself cannot account for the heavy Li isotope characteristics of the enriched MORB (Fig. 2). This is contrary to many popular notions of recycling (for example, refs 1, 24).

The heavy Li that is thought to be lost from the subducting oceanic crust^{9,10} is not evident in the Li isotope signatures of subduction zone volcanics⁷ and so it is thought to be transferred into the mantle through which slab-derived fluids pass, by isotopic reequilibration⁷. An appealing candidate for the heavy $\delta^7\text{Li}$, enriched material beneath the EPR is thus mantle that once overlaid the dehydrating slab⁴. A similar model has recently been invoked to account for other chemical signatures of enriched MORB^{25,26} and is consistent with the generally elevated $\delta^{18}\text{O}$ of such basalts in a high precision, global survey²⁷. Oxygen isotopes, however, do not clearly distinguish between subduction-modified mantle and subducted oceanic crust and it was previously argued that a contribution of recycled crust was responsible for the heavy $\delta^{18}\text{O}$ values measured in enriched MORB²⁷.

The complexity and range of the subduction processes make it difficult to model in detail the composition of subduction-modified mantle. As an illustrative calculation, however, we have used the composition of an amphibole peridotite (BZ 72) which has been analysed for Li isotopes and has been argued⁸ to record subduction zone enrichment. Not unexpectedly, the mixing curves of depleted mantle with this single amphibole peridotite (Fig. 2b–e) do not perfectly mimic the arrays of the EPR samples, but the general trends of Li isotope ratios with key tracers are well reproduced. Notably, mixing of the amphibole peridotite with depleted mantle can account for the decreasing La/Nb with increasing $\delta^7\text{Li}$ in our samples (Fig. 2c). It might seem counter-intuitive that our enriched component should have low La/Nb, as arc lavas typically have high La/Nb and, by inference, so will slab-derived fluids or melts that invade the mantle at subduction zones. Yet solid–fluid partitioning, in addition to the fluid composition itself, will influence the nature of sub-arc mantle enrichment. In particular, the growth of hydrous mineral phases might change the behaviour of elements normally considered

to be highly incompatible. In our modelled scenario, the presence of amphibole in BZ72 provides a significant host for Nb. Variable conditions of mineral stability between and within subduction zones, as well as different fluid compositions, can potentially lead to a variety of elemental enrichment patterns in subduction-modified mantle.

Subducting, mafic oceanic crust has radiogenic $^{87}\text{Sr}/^{86}\text{Sr}$ as a result of seafloor alteration and so the elevated $^{87}\text{Sr}/^{86}\text{Sr}$ of the enriched component (Fig. 2d) is again consistent with the isotopic exchange of slab-derived fluids with the overlying mantle. The Pb isotopes of our EPR samples^{11,12} show no systematic variations and are unremarkable for MORB (not shown), suggesting there is no sedimentary component in the subduction-enriched mantle identified here. Therefore, the unradiogenic Nd isotope signature of the enriched component is not an instantaneous product of subduction enrichment, but a consequence of subsequent evolution. At face value, the correlation of Sm/Nd with ε_{Nd} for the EPR basalts (Fig. 2f) could imply that an appropriate endmember was created by enrichment of a portion of depleted mantle at ~ 400 Myr ago. Other researchers²⁵ have noted similar apparent timescales between enrichment and magmatic sampling in Atlantic MORB, which they argued to represent residence times of enriched material in a continuously mixing upper mantle rather than discrete events.

From a physical perspective it is appealing that subduction-modified mantle is responsible for heterogeneities in the upper mantle. Subduction-modified mantle, initially chilled and viscously coupled to the down-going plate will warm, spall away and remix with ambient upper mantle. In contrast, the oceanic crust itself probably remains denser than the surrounding mantle and can potentially continue to sink even after it has become thermally equilibrated²⁸. Subduction-modified mantle is thus naturally mixed back into the upper mantle more efficiently than the subducted crust. The fate of the oceanic crust itself thus remains enigmatic³.

Received 10 February; accepted 1 August 2006.

- Allège, C. J. & Turcotte, D. L. Implications of a 2-component marble-cake mantle. *Nature* **323**, 123–127 (1986).
- Stracke, A., Bizimis, M. & Salters, V. J. M. Recycling oceanic crust: quantitative constraints. *Geochem. Geophys. Geosyst.* **4**, 8003, doi:10.1029/2001GC000223 (2003).
- Niu, Y. & O'Hara, M. J. Origin of ocean island basalts: a new perspective from petrology, geochemistry and mineral physics considerations. *J. Geophys. Res.* **108**, 2209, doi:10.1029/2002JB002048 (2003).
- Elliott, T., Jeffcoate, A. B. & Bouman, C. The terrestrial Li isotope cycle: light-weight constraints on mantle convection. *Earth Planet. Sci. Lett.* **220**, 231–245 (2004).
- Chan, L. H., Edmond, J. M., Thompson, G. & Gillis, K. Lithium isotopic composition of submarine basalts: implications for the lithium cycle in the oceans. *Earth Planet. Sci. Lett.* **108**, 151–160 (1992).

6. Decitre, S. *et al.* Behavior of Li and its isotopes during serpentinization of oceanic peridotites. *Geochim. Geophys. Geosyst.* **3**, 1007, doi:10.1029/2001GC000178 (2002).
7. Tomascak, P. B., Widom, E., Benton, L. D., Goldstein, S. L. & Ryan, J. G. The control of lithium budgets in island arcs. *Earth Planet. Sci. Lett.* **196**, 227–238 (2002).
8. Brooker, R. A., James, R. H. & Blundy, J. D. Trace elements and Li isotope systematics in Zabargad peridotites: evidence of ancient subduction processes in the Red Sea mantle. *Chem. Geol.* **212**, 179–204 (2004).
9. Zack, T., Tomascak, P. B., Rudnick, R. L., Dalpé, C. & McDonough, W. F. Extremely light Li in orogenic eclogites: the role of isotope fractionation during dehydration in subducted oceanic crust. *Earth Planet. Sci. Lett.* **208**, 279–290 (2003).
10. Wunder, B., Meixner, A., Romer, R. L. & Heinrich, W. Temperature-dependent isotopic fractionation of lithium between clinopyroxene and high-pressure hydrous fluids. *Contrib. Mineral. Petrol.* **151**, 112–120 (2006).
11. Niu, Y. L., Collerson, K. D., Batiza, R., Wendt, J. I. & Regelous, M. Origin of enriched-type mid-ocean ridge basalt at ridges far from mantle plumes: the East Pacific Rise at 11°20' N. *J. Geophys. Res.* **104**, 7067–7087 (1999).
12. Regelous, M. *et al.* Variations in the geochemistry of magmatism on the East Pacific Rise at 10°30' N since 800 ka. *Earth Planet. Sci. Lett.* **168**, 45–63 (1999).
13. Prinzhofer, A., Lewin, E. & Allègre, C. J. Stochastic melting of the marble cake mantle: evidence from local study of the East Pacific Rise at 12°50' N. *Earth Planet. Sci. Lett.* **92**, 189–206 (1989).
14. Jeffcoate, A. B., Elliott, T., Thomas, A. & Bouman, C. Precise, small sample size determination of lithium isotopic compositions of geological reference materials and modern seawater by MC-ICP-MS. *Geostand. Geoanal. Res.* **28**, 161–172 (2004).
15. Moriguti, T. & Nakamura, E. Across-arc variation of Li isotopes in lavas and implications for crust/mantle recycling at subduction zones. *Earth Planet. Sci. Lett.* **163**, 167–174 (1998).
16. Tomascak, P. B. & Langmuir, C. H. Lithium isotope variability in MORB. *Eos* **80**, F1086–F1087 (1999).
17. Seitz, H. M., Brey, G. P., Lahaye, Y., Durali, S. & Weyer, S. Lithium isotopic signatures of peridotite xenoliths and isotopic fractionation at high temperature between olivine and pyroxenes. *Chem. Geol.* **212**, 163–177 (2004).
18. Jeffcoate, A. B. *et al.* Li isotope fractionation in peridotites and mafic melts. *Geochim. Cosmochim. Acta* (in the press).
19. Magna, T., Wiechert, U. & Halliday, A. N. New constraints on the lithium isotope compositions of the Moon and terrestrial planets. *Earth Planet. Sci. Lett.* **243**, 336–353 (2006).
20. Chan, L. H., Alt, J. C. & Teagle, D. A. H. Lithium and lithium isotope profiles through the upper oceanic crust: a study of seawater-basalt exchange at ODP Sites 504B and 896A. *Earth Planet. Sci. Lett.* **201**, 187–201 (2002).
21. Gillis, K. M. & Coogan, L. A. Anatectic migmatites from the roof of an ocean ridge magma chamber. *J. Petrol.* **43**, 2075–2095 (2002).
22. Tomascak, P. B., Tera, F., Helz, R. T. & Walker, R. J. The absence of lithium isotope fractionation during basalt differentiation: new measurements by multicollector sector ICP-MS. *Geochim. Cosmochim. Acta* **63**, 907–910 (1999).
23. Lundstrom, C. C., Chaussidon, M., Hsui, A. T., Kelemen, P. & Zimmerman, M. Observations of Li isotopic variations in the Trinity Ophiolite: evidence for isotopic fractionation by diffusion during mantle melting. *Geochim. Cosmochim. Acta* **69**, 735–751 (2005).
24. Zindler, A., Staudigel, H. & Batiza, R. Isotope and trace element geochemistry of young Pacific seamounts: implications for the scale of upper mantle heterogeneity. *Earth Planet. Sci. Lett.* **70**, 175–195 (1984).
25. Donnelly, K. E., Goldstein, S. L., Langmuir, C. H. & Spiegelman, M. The origin of enriched ocean ridge basalts and implications for mantle dynamics. *Earth Planet. Sci. Lett.* **226**, 347–366 (2004).
26. Cooper, K. M., Eiler, J. M., Asimow, P. D. & Langmuir, C. H. Oxygen isotope evidence for the origin of enriched mantle beneath the mid-Atlantic ridge. *Earth Planet. Sci. Lett.* **220**, 297–316 (2004).
27. Eiler, J. M., Schiano, P., Kitchen, N. & Stolper, E. M. Oxygen-isotope evidence for recycled crust in the sources of mid-ocean-ridge basalts. *Nature* **403**, 530–534 (2000).
28. Hirose, K., Takafuji, N., Sata, N. & Ohishi, Y. Phase transition and density of subducted MORB crust in the lower mantle. *Earth Planet. Sci. Lett.* **237**, 239–251 (2005).
29. Sun, S.-S. & McDonough, W. F. in *Magmatism in the Ocean Basins* (eds Saunders, A. D. & Norry, M. J.) 313–345 (Blackwell, London, 1989).
30. Brueckner, H. K., Zindler, A., Seyler, M. & Bonatti, E. Zabargad and the isotopic evolution of the sub-Red Sea mantle and crust. *Tectonophysics* **150**, 163–176 (1988).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements Analytical work was supported by a Philip Leverhulme Prize awarded to T.E. A.J. was supported by an NERC studentship. Reviews by B. Leeman were appreciated. We also thank J. Blundy, C. Hawkesworth and D. Vance for comments on versions of the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to T.E. (tim.elliott@bris.ac.uk).

Reassembly of shattered chromosomes in *Deinococcus radiodurans*

Ksenija Zahradka^{1,2}, Dea Slade¹, Adriana Bailone³, Suzanne Sommer³, Dietrich Aeverbeck⁴, Mirjana Petranovic², Ariel B. Lindner¹ & Miroslav Radman^{1,5}

Dehydration or desiccation is one of the most frequent and severe challenges to living cells¹. The bacterium *Deinococcus radiodurans* is the best known extremophile among the few organisms that can survive extremely high exposures to desiccation and ionizing radiation, which shatter its genome into hundreds of short DNA fragments^{2–5}. Remarkably, these fragments are readily reassembled into a functional 3.28-megabase genome. Here we describe the relevant two-stage DNA repair process, which involves a previously unknown molecular mechanism for fragment reassembly called ‘extended synthesis-dependent strand annealing’ (ESDSA), followed and completed by crossovers. At least two genome copies and random DNA breakage are requirements for effective ESDSA. In ESDSA, chromosomal fragments with overlapping homologies are used both as primers and as templates for massive synthesis of complementary single strands, as occurs in a single-round multiplex polymerase chain reaction. This synthesis depends on DNA polymerase I and incorporates more nucleotides than does normal replication in intact cells. Newly synthesized complementary single-stranded extensions become ‘sticky ends’ that anneal with high precision, joining together contiguous DNA fragments into long, linear, double-stranded intermediates. These intermediates require RecA-dependent crossovers to mature into circular chromosomes that comprise double-stranded patchworks of numerous DNA blocks synthesized before radiation, connected by DNA blocks synthesized after radiation.

Both desiccation and radiation cause DNA double-strand breaks (DSBs)⁴, the most severe form of genomic damage^{6,7}. *Deinococcus radiodurans* survives exposures to ionizing radiation that produces over a thousand DSBs in each cell containing at least two genome copies, owing to a DNA repair process that accomplishes an efficient and precise reassembly of hundreds of short DNA fragments^{2,3,8}. The mechanism underlying this molecular transaction, formally akin to the computer-assisted contig assembly of shotgun-sequenced random genomic fragments, has remained a mystery for 50 yr (refs 9–15).

At least six mechanisms are known that could, either alone or in some combination, rejoin hundreds of partially overlapping chromosomal fragments (see Supplementary Information): non-homologous end-joining; homologous recombination at the fragment ends; intra- and interchromosomal single-strand annealing (SSA); synthesis-dependent-strand annealing (SDSA); break-induced replication; and copy choice (the switching of DNA replication from fragment to fragment). None of these mechanisms has been previously ruled out, but all are excluded by this study.

The requirement or involvement of DNA synthesis can be diagnostic

for the above DNA repair mechanisms (Supplementary Information). We therefore measured, in parallel, the kinetics of DNA fragment joining by pulsed-field gel electrophoresis (PFGE) and the rate of DNA synthesis by incorporation of [³H]thymidine during a 15-min pulse. After applying 7 kGy of γ -radiation, which shattered chromosomal DNA to fragments of about 20–30 kb (Fig. 1a), we observed a coincidence of DNA fragment assembly and massive DNA synthesis (Fig. 1a, b). This repair synthesis occurred before cell division at a rate much higher than that of DNA replication in the growing unirradiated cell cultures (Fig. 1b), and was absent in a *polA* strain (Fig. 1d) that showed no evidence of DNA repair (Fig. 1c). Neither repair nor synthesis was visible for 1.5 h after irradiation but, in the following 1.5 h, chromosomal DNA appeared fully reassembled (Fig. 1a, b) at 80–90% survival.

This largely RecA-independent fragment assembly¹⁶ (Fig. 1e) was accompanied by substantial DNA synthesis (Fig. 1f), but it did not produce complete chromosomes even after 24 h (Fig. 1e). These experiments establish a correlation between PolA-dependent (but not RecA-dependent) DNA synthesis and the reassembly of shattered *D. radiodurans* chromosomes that favours repair mechanisms that may (SSA and SDSA) or must (break-induced replication and copy choice) involve extensive DNA synthesis, rather than non-homologous end-joining and homologous recombination. RecA is clearly required for the appearance of full-size chromosomes (Fig. 1e), defining its key role in the late-stage process of maturation of circular chromosomes (see below).

How can small DNA fragments (Fig. 1a) sustain a rate of DNA synthesis that is much higher than that in normal DNA replication (Fig. 1b), and why is such synthesis correlated with fragment assembly (Fig. 1c, d)? To address these questions, the repair of the *D. radiodurans* DNA shattered by ionizing radiation was analysed by an adaptation of the classical Meselson–Stahl experiment¹⁷—that is, by DNA density labelling and analysis of its buoyant density by ultracentrifugation in caesium chloride equilibrium density gradients. The extent and pattern of DNA repair synthesis was monitored by the incorporation of a heavy analogue of thymidine, 5-bromo-deoxyuridine (BrdU). Whereas DNA replication in unirradiated *D. radiodurans* was semiconservative, the repair replication in irradiated cells generated a ‘distributive’¹⁷ pattern of DNA replication (Fig. 4 in Supplementary Data). The repaired chromosomes were fine patchworks (at both the double-strand and single-strand level) of DNA fragments synthesized before irradiation, interconnected by DNA blocks synthesized after irradiation. The old and new DNA blocks in the repaired chromosomes were of comparable size (~20–30 kb).

Unlike natural bases, BrdU is highly photosensitive; therefore, the

¹Université de Paris-Descartes, Faculté de Médecine, INSERM Site Necker, U571, 156 rue de Vaugirard, 75015 Paris, France. ²Division of Molecular Biology, Ruder Boskovic Institute, PO Box 180, 10002 Zagreb, Croatia. ³Institut de Génétique et Microbiologie, CNRS UMR8621, CEA LRC42V, Bâtiment 409, Université Paris-Sud, 91405 Orsay Cedex, France. ⁴Institut Curie, Section Recherche, UMR 2027 CNRS, Centre Universitaire de Paris-Sud, Bâtiment 110, 91405 Orsay Cedex, France. ⁵Mediterranean Institute for Life Sciences, Mestrovicova setaliste bb, 21000 Split, Croatia.

postulated old-Thy/new-BrdU patchwork structure of the repaired chromosomes could be tested by intracellular DNA photolysis. Only BrdU-substituted strands are sensitized to ultraviolet light, such that one-strand substitution in DNA duplex causes single-strand breaks, and two-strand substitution causes both single- and double-strand breaks^{18,19}. Whereas one-strand BrdU substitution of unirradiated *D. radiodurans* chromosomes did not significantly sensitize DNA to double-strand breakage by ultraviolet photolysis (data not shown), two-strand BrdU substitution resulted in extensive DNA photolysis (Fig. 2b). In γ -irradiated cells, ultraviolet photolysis degraded DNA by double-strand breakage, only when the DNA was repaired in BrdU medium, to fragment sizes similar to those seen immediately after γ -irradiation (Fig. 2a). As predicted by the ultraviolet insensitivity of old-Thy DNA, the ultraviolet-induced fragmentation of γ -irradiated chromosomal DNA repaired in BrdU medium did not increase beyond a saturating dose of $250\text{--}500\text{ J m}^{-2}$, whereas it did for

fully BrdU-substituted DNA (Fig. 2b). The pattern of ultraviolet photolysis suggested that most, perhaps all, reassembled DNA fragments were linked together by double-stranded blocks of newly synthesized DNA. The efficiency of ultraviolet photolysis suggested the presence of large sizes of newly synthesized double-stranded DNA patches in reconstituted chromosomes (see Supplementary Information).

To test the possibility that *D. radiodurans* DNA repair involves initial extensive synthesis of single-stranded DNA, as in SDSA-related mechanisms, we developed an immunofluorescence microscopy method to detect and measure, in single cells, the newly synthesized single- and double-stranded DNA. The method is based on the specificity of monoclonal antibodies to BrdU that bind the BrdU moiety in single- but not in double-stranded DNA. Similar to the experiment in Fig. 1b, newly synthesized DNA was labelled by 10-min BrdU pulses. Under native conditions, which revealed only the newly synthesized single-stranded DNA, unirradiated cells showed a low fluorescent background (Fig. 3a, c). After DNA denaturation, the antibody detected all newly synthesized DNA, and the intensity of fluorescent foci reflected the amount of BrdU incorporation during normal DNA replication (Fig. 3b, d).

In 7-kGy γ -irradiated cell cultures, under native conditions, BrdU pulses at different time periods caused the emergence of fluorescent foci only after a 1.5-h incubation corresponding to the onset of global DNA synthesis (Fig. 1b). Their frequency and intensity peaked at 3 h (Fig. 3a, c), and then both started to decrease and the foci disappeared by 6 h (Fig. 3a). Under denaturing conditions (Fig. 3b, d), the foci peaked at the same time as under native conditions, after which the rate started decreasing to that of normal DNA replication (see Fig. 1b). At the peak intensity, the fluorescence reflecting newly synthesized single-stranded DNA was about 15% of that reflecting all newly synthesized DNA. Thus, a considerable fraction, perhaps all, of

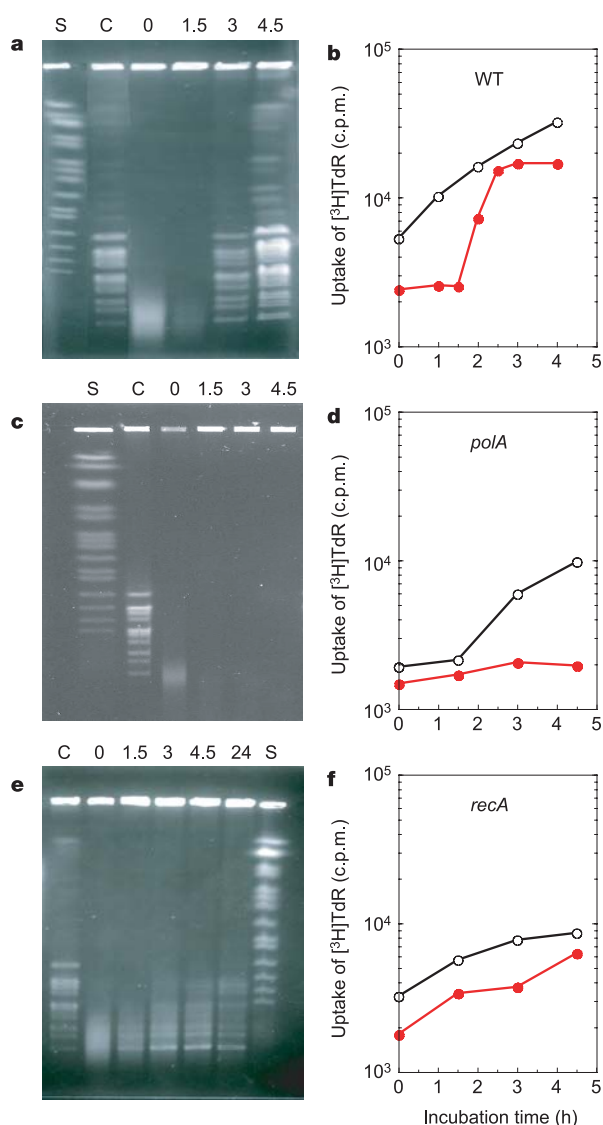


Figure 1 | DNA repair and synthesis after 7-kGy γ -irradiation of *D. radiodurans*. **a, c, e,** Kinetics of DNA repair in wild-type, *polA* and *recA* strains, respectively. 'C' indicates PFGE *NotI* restriction pattern of DNA from unirradiated cells; '0' indicates cells immediately after irradiation; other numbers indicate duration of cell growth after irradiation; 'S' indicates *S. cerevisiae* chromosomes used as molecular mass standards. **b, d, f,** Respective rates of DNA synthesis. Incorporation of [³H]thymidine during 15-min pulse labelling measures the global rate of DNA synthesis in irradiated (filled symbols) and unirradiated (open symbols) cultures.

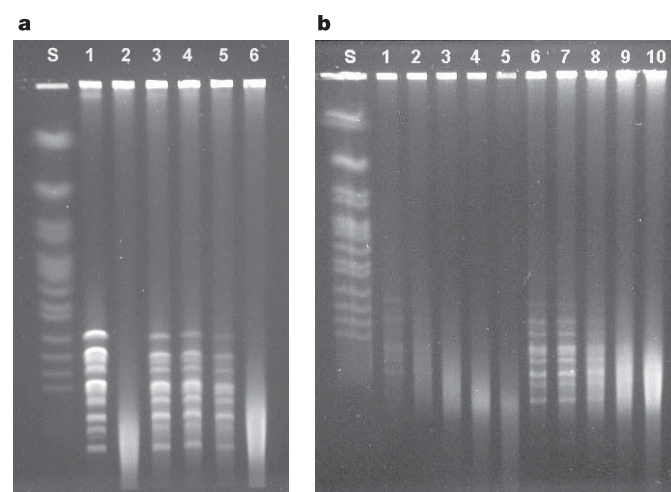


Figure 2 | Photolytic fragmentation of shattered *D. radiodurans* chromosomes reassembled in the presence of BrdU. **a,** PFGE *NotI* restriction pattern of *D. radiodurans* DNA (lanes 1–6) and *S. cerevisiae* molecular mass standards (S). Shown is DNA from unirradiated cells (lane 1), irradiated cells (lane 2), irradiated cells incubated for 3 h in TGY medium before (lane 3) or after (lane 4) exposure to $1,000\text{ J m}^{-2}$ of ultraviolet light, and irradiated cells repaired in BrdU-supplemented TGY medium before (lane 5) or after (lane 6) exposure to $1,000\text{ J m}^{-2}$ of ultraviolet light. **b,** Unlimited photolysis of fully BrdU-substituted DNA (lanes 1–5) versus limited photolysis of DNA repaired in BrdU (lanes 6–10). Shown is DNA from cells grown for 4.5 generations in BrdU-supplemented TGY medium before (lane 1) or after exposure to 100 J m^{-2} (lane 2), 250 J m^{-2} (lane 3), 500 J m^{-2} (lane 4) or $1,000\text{ J m}^{-2}$ (lane 5), and from irradiated cells repaired in BrdU-supplemented TGY medium (6) and then exposed to 100 J m^{-2} (lane 7), 250 J m^{-2} (lane 8), 500 J m^{-2} (lane 9) or $1,000\text{ J m}^{-2}$ (lane 10) of ultraviolet light.

DNA synthesis triggered by DNA fragmentation initially produced single-stranded DNA that was rapidly converted to double-stranded DNA.

We interpret our results as evidence that the capacity of *D. radiodurans* to reassemble hundreds of DNA fragments into intact circular chromosomes is due to a two-stage process (Fig. 4). The first stage involves a PolA-dependent mechanism that accomplishes most of fragment reassembly and that we call 'extended synthesis-dependent strand annealing' (ESDSA; Fig. 4a), as suggested by Meselson. The second, late-stage process of maturation of circular chromosomes seems to involve RecA-dependent crossovers (Fig. 4b). ESDSA differs from the standard 'limited' SDSA as described in the legend to Fig. 4. The apparent rapidity by which single strands are converted to double strands (Fig. 3a, b) suggests that the synthesis of complementary single strands often coincides in space and time. Such a coincidence can occur when two priming fragments, separated by a large gap, are bridged by a third fragment sharing overlapping homology and acting as a template for synthesis of the missing sequence. If the sequences of the two priming fragments were ABCD and GHIJ, and that of the bridging template fragment were DEFG, then the complete sequence of the assembled contigs would be ABCDEFGHIJ with the newly synthesized EF sequence.

The distinct features of the ESDSA model are that, first, it requires at least two chromosomal copies that are broken at different positions, and second; it involves a single-round multiplex polymerase chain reaction (PCR)-like step (Fig. 4, steps 2 and 3), resulting in long, newly synthesized, single-stranded overhangs that facilitate accurate annealing (Fig. 4, steps 4 and 5). The DNA repair pattern seen in PFGE (Fig. 1a, e) and in optical mapping experiments²⁰ shows that most DNA fragments 'grow' progressively and that most are used up in the chromosomal assembly process (Fig. 1). One can therefore imagine ESDSA 'chain reactions' involving numerous 'nucleation'

points in each cell producing growing linear double-stranded DNA repair intermediates²⁰. To become full-size chromosomes, such long, linear DNAs, either shorter or longer than the respective chromosome, require RecA-dependent intermolecular or intramolecular crossovers to produce unit-size circular chromosomes (Figs 1e and 4b). Although the *polA* mutant is deficient in ESDSA repair (Fig. 1), we cannot exclude the possibility that PolA initiates a DNA polymerase-III-catalysed single-strand elongation or contributes only to the maintenance of fragments undergoing PolA-dependent base excision repair of the damage caused by oxygen free radicals^{21,22}. Many details of the ESDSA mechanism have not been clarified, such as the priming step in DNA strand elongation, but the principal alternatives have been ruled out (see Supplementary Information).

Can ESDSA account for the apparent fidelity of DNA contig assembly in *D. radiodurans*? Generally, the larger the number of DNA fragments, the higher the precision required for avoiding their incorrect assembly and the longer the homology required. If there

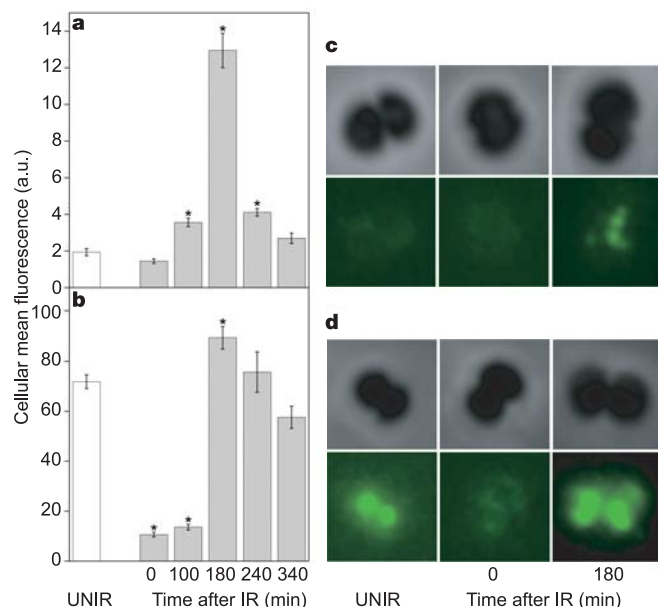


Figure 3 | Single-cell detection of repair-associated newly synthesized single-stranded DNA by immunofluorescence microscopy. *D. radiodurans* *thy⁻* unirradiated (UNIR) and irradiated (IR; 7 kGy) cells were pulse-labelled with BrdU at different time points and examined by immunofluorescence under native (**a**, **c**) and DNA denaturing (**b**, **d**) conditions. Unirradiated cells were labelled in the exponential growth phase. Error bars represent the s.e.m. Asterisks denote a statistically significant difference when compared with unirradiated exponential cells (*t*-test, $P < 10^{-2}$). **c**, **d**, Representative images (fluorescence images adjusted to the same intensity scale and phase-contrast images) of cells before (UNIR) and after (IR) irradiation.

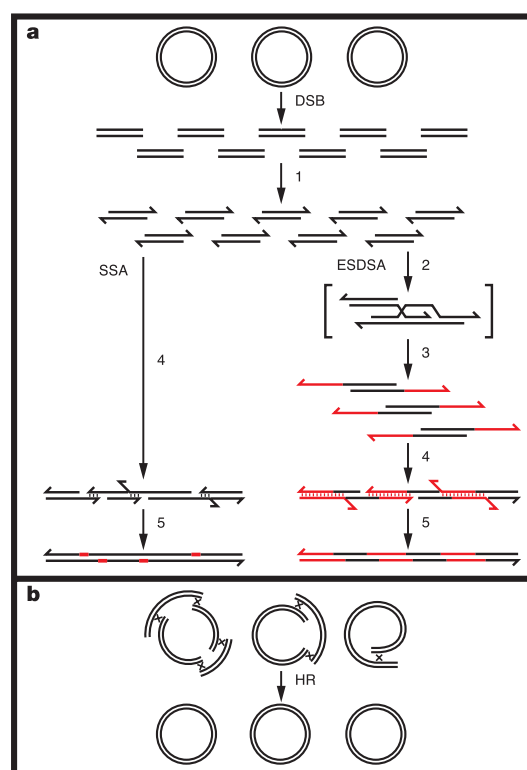


Figure 4 | Two-stage process of reconstitution of shattered chromosomes. **a**, Two alternative strand annealing models for reassembly of small chromosomal fragments into long linear intermediates. Several genomic copies of *D. radiodurans* undergo random DNA double-strand breakage, producing numerous fragments. After 5'-to-3' end resection (step 1), DNA fragments can rejoin directly through the annealing of complementary single-stranded overhangs of overlapping fragments belonging to different chromosomal copies (step 4, interchromosomal SSA pathway); the gaps are repaired by synthesis and strand excess is trimmed by nucleases (step 5). In the ESDSA pathway, the end-resected fragments (step 1) prime synthesis using the homologous regions of partially overlapping fragments as a template (step 2), presumably through a moving D-loop (bracketed intermediate). The strand extension can run to the end of the template, producing fragments with long, newly synthesized (red) single-stranded overhangs (step 3) that can *a priori* engage in several rounds of extension until they find a complementary partner strand (steps 4 and 5 are the same in ESDSA and SSA). Hydrogen bonds are indicated only for interfragment associations. **b**, To mature into unit-size circular chromosomes, the overlapping, long, linear fragments, or linear intermediates longer than the chromosome, require crossovers (X) by means of RecA-dependent homologous recombination (HR).

is an unusually large number of DSBs, then an unusually long homology, and thus single-strand exposure, is required for an accurate reassembly. Unusually long overhangs are newly synthesized in the course of ESDSA (Fig. 4a). However, all types of homology-based fragment reassembly raise the issue of accuracy when DNA fragments contain repetitive sequences. This accuracy could be maintained if the tails of the fragments were much longer than the longest repetitive sequences. Annealing only a limited repeated sequence block within two long non-complementary single-stranded overhangs could not readily link the two fragments. Aberrant DNA associations through partially annealed structures are promptly dissociated by DNA helicases. Our experiments suggest that the size of newly synthesized overhangs is about 20–30 kb, which is much longer than the longest repetitive sequences (insertion sequences of ~1 kb)^{9,23} in *D. radiodurans* (see also Supplementary Information).

A high-fidelity ESDSA process requires that all overhangs of a fragment must be extended by copying fragments that are contiguous in the intact chromosome. We can think of four strategies for assuring the fidelity of both synthesis priming and strand annealing: first, homologous pairing of recessed double-stranded DNA fragments before the initiation of D-loops, perhaps by the peculiar deinococcal RecA¹²; second, editing the pairing process by mismatch repair proteins^{24,25}; third, repeat-binding proteins⁹ preventing sequence repeats from becoming single stranded or from annealing; and fourth, stable secondary structures (hairpins) of repetitive sequences²³ preventing annealing with the same repeat in a non-complementary single-stranded partner.

The robustness of *D. radiodurans* has presumably evolved under harsh non-reproductive conditions such as those present in arid desert environments²⁶. Extreme desiccation accompanied by extensive DNA breakage⁴, as in our irradiation experiments, leads to the 'clinical death' of deinococcal cells. On supply of water and ions, however, the dead bacterium 'resurrects' by reconstituting its shattered genome²⁷. *D. radiodurans* can be seen as a bacterial model of long-lived non-dividing neurons. Thus, exploring mechanisms of deinococcal robustness could inspire approaches in anti-ageing research and regenerative medicine. Reproducing *in vitro* the DNA recombination repair of *D. radiodurans* under low-fidelity conditions could provide a tool for the shuffling of genomic fragments from the whole biosphere.

METHODS

See Supplementary Information for a full description of the methods used.

Strains. *D. radiodurans* strains R1 wild type²⁸ and a *thy*[−] derivative (this work), GY10922 $\Delta(cinA-recA)::kan^r$ and IRS501 *polA* (J. R. Battista) were used. Bacteria were grown in TGY medium at 30°C to the late exponential phase and exposed to 7 kGy γ -irradiation, resulting in 80–90% survival of the wild-type strain.

Measurement of DNA size, density and synthesis. The kinetics of DNA fragment joining was measured by PFGE. Irradiated cells were embedded in agarose plugs, lysed and treated with *NotI* restriction enzyme⁴. The plugs were subjected to pulsed electric field in 0.5× TBE buffer using a CHEF-DR III electrophoresis system (Bio-Rad) at 6 V cm^{−2} for 20 h at 14°C, with a linear pulse ramp of 50–90 s and a switching angle of 120°.

The rate of DNA synthesis was measured by pulse labelling 0.5-ml samples of exponential cultures with [³H]thymidine for 15 min. Pulses were terminated by addition of ice-cold 10% trichloroacetic acid and the precipitated counts measured. For DNA density gradient analysis, *D. radiodurans thy*[−] cells were radioactively and density labelled during growth in the presence of [³H]thymidine and BrdU, respectively. The DNA was isolated and its buoyant density was analysed by ultracentrifugation in CsCl density gradients.

Intracellular photolysis of BrdU-substituted DNA. For DNA photolysis, irradiated *thy*[−] cells were grown in BrdU-supplemented TGY medium for 3 h, then starved in buffer, ice cooled, exposed in a thin layer to the indicated doses of 254-nm ultraviolet light and embedded in agarose plugs for PFGE analysis of DNA.

Detection of single-strand DNA synthesis at the single cell level. The newly synthesized DNA was detected and measured by immunofluorescence

microscopy. Exponentially grown *thy*[−] cell suspensions were exposed to 10-min BrdU pulses, fixed in methanol, rendered permeable for antibodies, and then treated with primary antibodies to BrdU, followed by fluorescence-tagged secondary antibodies. Fluorescence was analysed by microscopy. Image analysis was done with Metamorph software.

Received 24 May; accepted 9 August 2006.

Published online 27 September 2006.

- Potts, M. Desiccation tolerance: a simple process? *Trends Microbiol.* **9**, 553–559 (2001).
- Minton, K. W. DNA repair in the extremely radioresistant bacterium *Deinococcus radiodurans*. *Mol. Microbiol.* **13**, 9–15 (1994).
- Battista, J. R., Earl, A. M. & Park, M. J. Why is *Deinococcus radiodurans* so resistant to ionizing radiation? *Trends Microbiol.* **7**, 362–365 (1999).
- Mattimore, V. & Battista, J. R. Radioresistance of *Deinococcus radiodurans*: functions necessary to survive ionizing radiation are also necessary to survive prolonged desiccation. *J. Bacteriol.* **178**, 633–637 (1996).
- Sanders, S. W. & Maxcy, R. B. Isolation of radiation-resistant bacteria without exposure to radiation. *Appl. Environ. Microbiol.* **38**, 436–439 (1979).
- Krasin, F. & Hutchinson, F. Repair of DNA double-strand breaks in *Escherichia coli* which requires *recA* function and the presence of duplicate genome. *J. Mol. Biol.* **116**, 81–98 (1977).
- Fujimori, A. et al. Rad52 partially substitutes for Rad51 paralog XRCC3 in maintaining chromosomal integrity in vertebrate cells. *EMBO J.* **20**, 5513–5520 (2001).
- Dean, C. J., Feldschreiber, P. & Lett, J. T. Repair of X-ray damage to the deoxyribonucleic acid in *Micrococcus radiodurans*. *Nature* **209**, 49–52 (1966).
- White, O. et al. Genome sequence of the radioresistant bacterium *Deinococcus radiodurans*. *Science* **286**, 1571–1577 (1999).
- Ferreira, A. C. et al. Characterization and radiation resistance of new isolates of *Rubrobacter radiotolerans* and *Rubrobacter xylanophilus*. *Extremophiles* **3**, 235–238 (1999).
- Narumi, I. Unlocking radiation resistance mechanisms: still a long way to go. *Trends Microbiol.* **11**, 422–425 (2003).
- Cox, M. M. & Battista, J. R. *Deinococcus radiodurans*—the consummate survivor. *Nature Rev. Microbiol.* **3**, 882–892 (2005).
- Harsojo, Kitayama, S. & Matsuyama, A. Genome multiplicity and radiation resistance in *Micrococcus radiodurans*. *J. Biochem.* **90**, 877–880 (1981).
- Levin-Zaidman, S. et al. Ringlike structure of the *Deinococcus radiodurans* genome: a key to radioresistance? *Science* **299**, 254–256 (2003).
- Zimmerman, J. M. & Battista, J. R. A ring-like nucleoid is not necessary for radioresistance in *Deinococcaceae*. *BMC Microbiol.* **5**, 17–27 (2005).
- Daly, M. J. & Minton, K. W. An alternative pathway of recombination of chromosomal fragments precedes *recA*-dependent recombination in the radioresistant bacterium *Deinococcus radiodurans*. *J. Bacteriol.* **178**, 4461–4471 (1996).
- Meselson, M. & Stahl, F. W. The replication of DNA. *Cold Spring Harb. Symp. Quant. Biol.* **23**, 9–12 (1958).
- Lion, M. B. Search for a mechanism for the increased sensitivity of 5-bromouracil-substituted DNA to ultraviolet radiation. II. Single-strand breaks in the DNA of irradiated 5-bromouracil-substituted T3 coliphage. *Biochim. Biophys. Acta* **209**, 24–33 (1970).
- Lohman, P. H. M., Bootsma, D. & Hey, A. H. The influence of 5-bromodeoxyuridine on the induction of breaks in deoxyribonucleic acid of cultivated human cells by X-irradiation and ultraviolet light. *Radiat. Res.* **52**, 627–641 (1972).
- Lin, J. et al. Whole-genome shotgun optical mapping of *Deinococcus radiodurans*. *Science* **285**, 1558–1562 (1999).
- Milligan, J. R. et al. DNA strand-break yield after incubation with base excision repair endonucleases implicate hydroxyl radical pairs in double-strand break formation. *Int. J. Radiat. Biol.* **76**, 1475–1483 (2000).
- Lindahl, T. DNA repair enzymes. *Annu. Rev. Biochem.* **51**, 61–87 (1982).
- Makarova, K. S. et al. Genome of the extremely radiation-resistant bacterium *Deinococcus radiodurans* viewed from the perspective of comparative genomics. *Microbiol. Mol. Biol. Rev.* **65**, 44–79 (2001).
- Rayssiguier, C., Thaler, D. S. & Radman, M. The barrier to recombination between *Escherichia coli* and *Salmonella typhimurium* is disrupted in mismatch-repair mutants. *Nature* **342**, 396–401 (1989).
- Mennecier, S., Coste, G., Servant, P., Bailone, A. & Sommer, S. Mismatch repair ensures fidelity of replication and recombination in the radioresistant organism *Deinococcus radiodurans*. *Mol. Genet. Genomics* **272**, 460–469 (2004).
- Rainey, F. A. et al. Extensive diversity of ionizing-radiation-resistant bacteria recovered from Sonoran Desert soil and description of nine new species of the genus *Deinococcus* from a single soil sample. *Appl. Environ. Microbiol.* **71**, 5225–5235 (2005).
- Harris, D. R. et al. Preserving genome integrity: the DdrA protein of *Deinococcus radiodurans* R1. *PLoS Biol.* **2**, 1629–1639 (2004).
- Anderson, A. W., Nordon, H. C., Cain, R. F., Parrish, G. & Duggan, D. Studies on a radio-resistant micrococcus. I. Isolation, morphology, cultural characteristics and resistance to γ radiation. *Food Technol.* **10**, 575–578 (1956).
- Bonacossa de Almeida, C., Coste, G., Sommer, S. & Bailone, A. Quantification

of RecA protein in *Deinococcus radiodurans* reveals involvement of RecA, but not LexA, in its regulation. *Mol. Genet. Genomics* **268**, 28–41 (2002).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank J. Battista for initial help with *Deinococcus* and for the *polA* mutant; M. Meselson, B. Wagner, A. Stark, D. Zahradka and members of the M.R. lab for reading this manuscript and/or providing advice; M. Blazevic for technical assistance on γ -irradiation; and Mikula Radman for illustrations. K.Z. held a fellowship from the Necker Institute during work in M.R.'s laboratory. D.S. holds a Boehringer Ingelheim Foundation predoctoral fellowship and A.B.L. is a Marie Curie fellow. The laboratory in the Université de Paris-Descartes was funded by INSERM and the Necker Institute (Mixis/PLIVA contract); that in the Institut de Génétique et Microbiologie by EDF-France and CNRS (GEOMEX);

that in the Institut Curie by EDF-France; and that in the Ruder Boskovic Institute by the Croatian Ministry of Science, Education and Sports.

Author Contributions Experiments in Figs 1 and 2 and Supplementary Fig. 4 were carried out by K.Z.; those in Fig. 3 by D.S. A.B.L. provided expertise in microscopy; and the global experimental design was by M.R. The results of Fig. 1b, f, were obtained in, and with the scientific and material support of, the Institut de Génétique et Microbiologie; those of Fig. 1a, e, in the Institut Curie; those of Figs 1c, d, 2 and Supplementary Fig. 4 in the Ruder Boskovic Institute, and those of Fig. 3 in the Université de Paris-Descartes.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to M.R. (radman@necker.fr).

LETTERS

Functional epistasis on a common MHC haplotype associated with multiple sclerosis

Jon W. Gregersen¹, Kamil R. Kranc², Xiayi Ke³, Pia Svendsen¹, Lars S. Madsen⁴, Allan Randrup Thomsen⁵, Lon R. Cardon³, John I. Bell^{2,6} & Lars Fugger^{1,2}

Genes in the major histocompatibility complex (MHC) encode proteins important in activating antigen-specific immune responses. Alleles at adjacent MHC loci are often in strong linkage disequilibrium; however, little is known about the mechanisms responsible for this linkage disequilibrium. Here we report that the human MHC HLA-DR2 haplotype, which predisposes to multiple sclerosis^{1–3}, shows more extensive linkage disequilibrium than other common caucasian HLA haplotypes in the DR region and thus seems likely to have been maintained through positive selection. Characterization of two multiple-sclerosis-associated HLA-DR alleles at separate loci by a functional assay in humanized mice indicates that the linkage disequilibrium between the two alleles may be due to a functional epistatic interaction, whereby one allele modifies the T-cell response activated by the second allele through activation-induced cell death. This functional epistasis is associated with a milder form of multiple-sclerosis-like disease. Such epistatic interaction might prove to be an important general mechanism for modifying exuberant immune responses that are deleterious to the host and could also help to explain the strong linkage disequilibrium in this and perhaps other HLA haplotypes.

The HLA-DR2 haplotype in the human MHC contains the DR alleles *DRB1*1501* (*DR2b*) and *DRB5*0101* (*DR2a*) at the respective *DRB1* and *DRB5* loci, located 85 kb apart. In all ethnic groups studied so far, these alleles show almost complete linkage disequilibrium with haplotype frequencies varying across populations^{4–10}. Previous data suggest that this haplotype is exceptional in the extent of its linkage disequilibrium¹¹. Using extensive single nucleotide polymorphism (SNP) genotyping data available for this genomic region¹¹, we analysed the extent and length of linkage disequilibrium in caucasian samples. We found that the HLA-DR2 haplotype shows the most extensive preservation relative to all other common HLA-DR haplotypes (Fig. 1a–c). These data thus suggest that strong positive selection is operating on this haplotype and that these DR loci may confer a particular selective advantage.

In theory, the inseparability of *DR2a* from *DR2b* could imply the persistence of a founder haplotype; however, its ubiquity in widely variant ethnic groups^{4–10} argues against this being the only explanation. Another possibility might be an advantageous epistatic interaction between these alleles. As in Bateson's original definition¹², we use the term 'epistasis' here to mean the masking of a functional effect of one allele at one locus by an allele at a separate locus¹³. If the *DR2b-DR2a* combination were very advantageous, it might favour much stronger linkage disequilibrium between these coding HLA-DR alleles than among interspersed or flanking anonymous SNPs. Studies have highlighted such differences between SNPs and HLA

alleles^{11,14}. Analysis of the available SNP data in the *DR2* region for Nigerian (Yoruba) and Asian samples in the International HapMap project¹⁵ suggests that there are population differences in linkage disequilibrium patterns for SNPs (Supplementary Fig. 1). This observation is in stark contrast to the strong linkage disequilibrium between *DR2b* and *DR2a*^{4–10}, and strongly suggests the co-selection of epistatic alleles.

In support of these genetic observations, we have obtained independent biological evidence for an epistatic interaction between

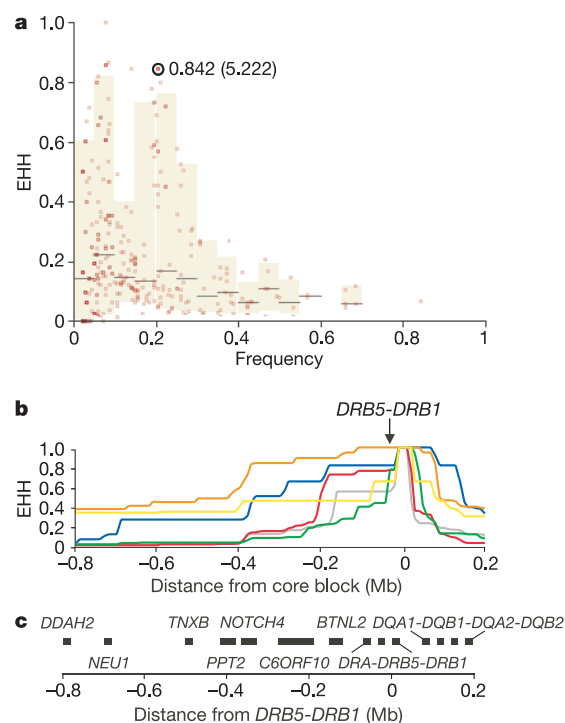


Figure 1 | Analysis of selection around HLA-DRB1 and HLA-DRB5 loci using extended haplotype homozygosity. **a**, Distribution of extended haplotype homozygosity (EHH) scores by frequency at 300 kb from the core block. The circled outlier is the HLA-DR2 haplotype with its EHH score and relative EHH score given in parentheses. The shaded bars are 5% and 95% percentiles. **b**, Plot of EHH. The frequencies of the haplotypes are 26% (red), 21% (orange, *DR2* outlier), 18% (green), 16% (grey), 8% (yellow) and 6% (blue). **c**, Approximate gene locations according to the UCSC Genome Bioinformatics website (<http://genome.ucsc.edu>; March 2006).

¹Department of Clinical Immunology, Aarhus University Hospital, Skejby Sygehus, 8200 N, Aarhus, Denmark. ²MRC Human Immunology Unit, Weatherall Institute of Molecular Medicine, John Radcliffe Hospital, University of Oxford, OX3 9DS, Oxford, UK. ³Wellcome Trust Centre for Human Genetics, University of Oxford, OX3 7BN, Oxford, UK.

⁴Department of Clinical Immunology, Copenhagen University Hospital, Rigshospitalet, 2100 Ø, Copenhagen, Denmark. ⁵Institute of Medical Microbiology and Immunology, University of Copenhagen, 2200 N, Copenhagen, Denmark. ⁶Office of the Regius Professor of Medicine, The Richard Doll Building, University of Oxford, OX3 7BN, Oxford, UK.

Table 1 | Incidence, severity and mean onset of EAE in humanized mice

Transgene	Disease incidence	Maximum severity score (mean $\pm$ s.d.)	Onset in days (mean $\pm$ s.d.)
<i>Hy</i> ⁺ - <i>DR2b</i> ⁺	38/44 (86.4%)	7.1 $\pm$ 1.0	181.9 $\pm$ 66.7
<i>Hy</i> ⁺ - <i>DR2b</i> ⁺ - <i>DR2a</i> ⁺	30/55* (54.5%)	6.0 $\pm$ 1.7**	221.6 $\pm$ 70.1***
<i>Ob</i> ⁺ - <i>DR2b</i> ⁺	9/11 (81.8%)	6.8 $\pm$ 1.3	88.1 $\pm$ 26.0
<i>Ob</i> ⁺ - <i>DR2b</i> ⁺ - <i>DR2a</i> ⁺	7/8 (87.5%)	6.3 $\pm$ 1.8	95.8 $\pm$ 51.0
<i>Hy</i> ⁺ - <i>DR2a</i> ⁺	0/30	0	-

* $P < 0.001$, ** $P < 0.003$ and *** $P < 0.022$ versus *Hy*⁺-*DR2b*⁺ transgenic mice.

DR2b and *DR2a*. Our functional studies in humanized mice have dissected the roles of these alleles in predisposition to multiple sclerosis. Previous genetic studies localized the dominant genetic determinant of susceptibility to the region around the *DRB1* and *DRB5* loci on the HLA-DR2 haplotype^{1,2}, but dissection of the precise DNA variant accounting for this susceptibility has been constrained by the almost complete linkage disequilibrium between the *DR2b* and *DR2a* alleles¹⁶. Transferring these genes into mice has enabled us to analyse their contributions, either separately or together, to the autoimmune T-cell response and to the resulting multiple-sclerosis-like disease experimental autoimmune encephalomyelitis (EAE).

We used two T-cell antigen receptors (TCRs) derived from individuals affected with multiple sclerosis: *Hy*.2E11 (*Hy*) and *Ob*.1A12 (*Ob*)^{17,18}. Both TCRs recognize the immunodominant epitope in the multiple-sclerosis-relevant autoantigen myelin basic protein, MBP(85–99), when presented by the *DR2b* molecule (encoded by *DRB1**1501 and the invariant *DRA1**0101 allele; Supplementary Fig. 2 and Supplementary Table 1). The *Hy* TCR also recognizes several different epitopes presented by the *DR2a* molecule¹⁹ (encoded by *DRB5**0101 and the invariant *DRA1**0101 allele), a cross-reactivity not seen with the *Ob* TCR¹⁹. We generated humanized transgenic mice expressing the *Hy* TCR or *Ob* TCR and *DR2b* or *DR2a* or both of these HLA class II alleles^{19,20}. Because they had been crossed to *Rag2*^{-/-} mice, which do not express endogenous murine TCRs, the only TCR expressed by the mice was the human transgene-encoded TCR.

The *Hy*⁺-*DR2b*⁺ mice spontaneously developed very severe EAE (Table 1). Addition of the *DR2a* transgene significantly decreased its incidence, reduced its severity and delayed its onset (Table 1). We also observed complete spontaneous clinical remissions in about 20% of *Hy*⁺-*DR2b*⁺-*DR2a*⁺ transgenic mice, whereas none of the 38 *Hy*⁺-*DR2b*⁺ transgenic mice recovered ($P \leq 0.014$) from EAE. In addition, spontaneous primary progressive EAE was very prevalent in *Ob*⁺-*DR2b*⁺ transgenic mice and was not affected by the introduction of *DR2a* (Table 1). These results indicate that the modifying effect of *DR2a* may not simply be due to a nonspecific event, such as DRA chain stealing, but may be due to an epistatic interaction with *DR2b* that depends on a cross-reactive TCR. This conclusion was corroborated by separate experiments showing that neither *CD4*⁺ T-cell frequencies nor the severity of EAE in *Hy*⁺ - *DR2b*⁺ transgenic mice was affected by the addition of another HLA-DR allele, *DRB1**0401 (data not shown). Taken together, our results imply that, at least in this setting, the *DR2b* allele mediates EAE, whereas *DR2a* functions as a genetic modifier of the resulting disease.

One possible explanation for our observations is that the *DR2a* molecules induce deletion of *Hy*⁺*CD4*⁺ T cells and thereby modulate the severity of the *DR2b*-dependent disease. We therefore compared the frequencies of *Hy*⁺*CD4*⁺ T cells in thymus and blood from *Hy*⁺-*DR2b*⁺ and *Hy*⁺-*DR2b*⁺-*DR2a*⁺ transgenic mice on a *Rag2*^{-/-} background. Whereas we observed no significant difference in their thymus, we found markedly lower frequencies of *Hy*⁺*CD4*⁺ T cells in blood from the *Hy*⁺-*DR2b*⁺-*DR2a*⁺ mice (Table 2). By contrast, we saw no such differences in the frequencies of *Ob*⁺*CD4*⁺ T cells between *Ob*⁺-*DR2b*⁺-*DR2a*⁺ and *Ob*⁺-*DR2b*⁺ transgenic mice (Table 2). These results suggest that the *DR2a* molecules do not induce central tolerance in *Hy*⁺-*DR2b*⁺-*DR2a*⁺ transgenic mice, but

Table 2 | Frequency of autoreactive T cells in humanized mice

Transgene	Organ	Median frequency of <i>Hy</i> ⁺ <i>CD4</i> ⁺ or <i>Ob</i> ⁺ <i>CD4</i> ⁺ T cells (25–75 percentile)
<i>Hy</i> ⁺ - <i>DR2b</i> ⁺	Blood	11.7 (1.18–32.7)
<i>Hy</i> ⁺ - <i>DR2b</i> ⁺ - <i>DR2a</i> ⁺	Blood	1.40 (0.06–9.0)*
<i>Hy</i> ⁺ - <i>DR2b</i> ⁺	Thymus	0.17 (0.09–0.45)
<i>Hy</i> ⁺ - <i>DR2b</i> ⁺ - <i>DR2a</i> ⁺	Thymus	0.34 (0.12–0.45)
<i>Ob</i> ⁺ - <i>DR2b</i> ⁺ - <i>DR2a</i> ⁺	Blood	1.40 (0.70–2.20)
<i>Ob</i> ⁺ - <i>DR2b</i> ⁺	Blood	2.05 (0.60–5.8)

* $P \leq 0.007$ versus *Hy*⁺*CD4*⁺ frequency in blood from *Hy*⁺-*DR2b*⁺ transgenic mice.

instead reduce the number of autoreactive *Hy*⁺*CD4*⁺ T cells in the periphery.

One mechanism of peripheral tolerance is clonal deletion of self-reactive T cells by activation-induced cell death (AICD)²¹. The hallmark of AICD is activation, followed by proliferative expansion and apoptotic cell death of the activated T cells. We tested whether *DR2a* molecules could induce peripheral deletion by AICD of adoptively transferred *Hy*⁺*CD4*⁺ T cells. In *DR2a*⁺ and *DR2b*⁺-*DR2a*⁺ recipients, the cells underwent multiple divisions within 5 days of transfer (Fig. 2a) and showed a *CD44*⁺*CD62L*⁻ phenotype (Supplementary Fig. 3) consistent with activation, whereas almost all cells remained undivided in *DR2b*⁺ recipients (Fig. 2a) and predominantly had a *CD44*⁻*CD62L*⁺ naive phenotype (Supplementary Fig. 3). Moreover, by day 5, the rate of apoptosis in *Hy*⁺*CD4*⁺ T cells adoptively transferred into *DR2b*⁺-*DR2a*⁺ and *DR2a*⁺ mice was markedly increased as compared with that in *DR2b*⁺ mice (Fig. 2b). *Hy*⁺*CD4*⁺ T cells were barely detectable in spleens and blood from the former groups after 15 d from transfer but still persisted in the latter group (Fig. 2c and Supplementary Fig. 4). These results indicate that *DR2a* molecules induce peripheral deletion of *Hy*⁺*CD4*⁺ T cells by means of AICD. Thus, they show epistasis between the *DR2b* and *DR2a* alleles at a functional level (Supplementary Fig. 5).

One of the hallmarks of the human MHC is the strong linkage disequilibrium that extends across many HLA haplotypes. Although

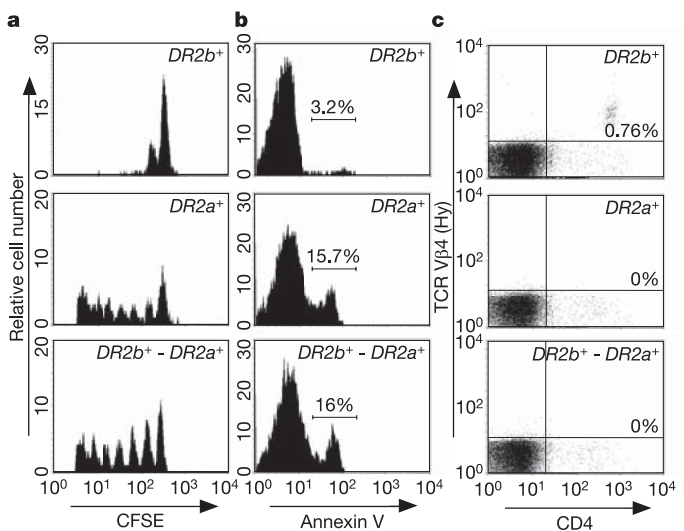


Figure 2 | *DR2a* induces peripheral deletion of *Hy*⁺*CD4*⁺ T cells by AICD. **a**, Proliferation of CFSE-labelled *Hy*⁺*CD4*⁺ T cells 5 d after transfer into *DR2b*⁺, *DR2a*⁺ and *DR2b*⁺-*DR2a*⁺ transgenic mice. Division of transferred T cells in spleens of the recipient mice was measured by CFSE dilution using flow cytometry. **b**, Apoptosis of *Hy*⁺*CD4*⁺ T cells 5 d after transfer into the indicated recipient mice. Apoptotic *Hy*⁺*CD4*⁺ T cells in spleens of the recipient mice were detected by Annexin-V staining. **c**, Frequency of *Hy*⁺*CD4*⁺ T cells in spleens of *DR2b*⁺, *DR2a*⁺ and *DR2b*⁺-*DR2a*⁺ recipient mice 15 d after transfer.

it was originally thought to be unique to the MHC, similar regions of strong linkage disequilibrium are now known to exist elsewhere in the genome¹⁵. In all cases, the mechanisms for their generation and maintenance are not clear. Several possible explanations exist for long-range linkage disequilibrium such as that apparent on the HLA-DR2 haplotype. The simplest is that strong positive selection for a component of the haplotype has expanded its frequency in recent evolutionary history, creating a founder haplotype that has had inadequate evolutionary time to degrade. This explanation, although potentially important in explaining much linkage disequilibrium in the HLA region, is unlikely to provide the only mechanism for these effects, however, because the *DR2b* and *DR2a* alleles are in strong linkage disequilibrium in all ethnic populations in which they have been studied^{4–10}. In addition, this tight linkage disequilibrium is not always paralleled by tight linkage disequilibrium between anonymous SNPs in this region across populations. These data indicate that recent selection of founder haplotypes is unlikely and that the *DRB* alleles on this haplotype probably have repeatedly re-associated through the strength of their functional interaction. Together, therefore, our functional and genetic data support the hypothesis that epistasis is at least partly responsible for this strong linkage disequilibrium.

The epistatic interaction between the *DR2a* and *DR2b* alleles requires a cross-reactive TCR restricted by both corresponding MHC class II molecules. This leads to a mechanism of immune modulation, whereby *DR2b* molecules mediate a strong immune response, whereas *DR2a* molecules modify this effect through peripheral T-cell deletion by means of apoptosis. Although the natural antigenic stimulus responsible for the co-selection through epistasis of the two DR alleles is not known, it is likely that in our experimental model a self-antigen produces similar epistatic functional effects. In our model, coexpression of the *DR2a* and *DR2b* molecules reduces the magnitude and severity of multiple-sclerosis-like disease and produces remissions that are characteristic of the most common form of multiple sclerosis seen in humans—relapsing-remitting disease³. The relevance of our data to the human situation is supported by the observation that intact apoptosis pathways are required for spontaneous clinical remissions in other relapsing remitting EAE models and are crucial for controlling disease severity²². Taken together, these observations suggest a situation whereby *DR2a* operates in concert with *DR2b* and partially deletes autoreactive T cells to limit autoreactive responses triggered by *DR2b*, thus creating a less severe and more clinically relevant phenotype (Supplementary Fig. 5).

Although this functional epistasis is occurring in an experimental model, its ability to modulate normal immune responses is apparent. Vigorous host immune responses are ultimately responsible for much of the pathology and tissue damage in various infectious diseases such as viral hepatitis²³ and schistosomiasis²⁴ in humans and viral lymphocytic choriomeningitis in mouse²⁵. Maintaining a balance between controlling the infectious pathogen and causing tissue damage requires a balance in the immune response. The epistasis that we have described provides modulation of the immune response activated by one HLA allele. This epistatic modulation of the severity of such a response could have an important general regulatory role and may therefore be more widespread throughout the MHC and might also contribute to patterns of linkage disequilibrium on other MHC haplotypes.

Separating the effects of individual human alleles at separate loci that show strong linkage disequilibrium can be extremely difficult. The experimental model that we have used for these studies has wide implications for those studying functional contributions of other genetic loci held together in strong linkage disequilibrium. Our humanized murine model has enabled us to characterize the independent and interactive role of two multiple-sclerosis-associated DR alleles *in vivo* and to demonstrate functional effects. By extension, this type of functional approach can provide a powerful mechanism

for determining the relevant role of other individual alleles that are indistinguishable in humans because of high or complete linkage disequilibrium.

METHODS

Extended haplotype homozygosity analysis. Extended haplotype homozygosity (EHH)²⁶ analysis was done with Sweep (<http://www.broad.mit.edu/mpg/sweep/>) using the Centre d'Etude du Polymorphisme Humain (CEPH) SNP data sets¹¹ as described in Supplementary Methods.

Integrated haplotype score analysis. We downloaded the genome-wide integrated haplotype scores²⁷ for the HapMap phase I CEU, CHB + JPT and YRI samples from the Pritchard Lab website (<http://pritch.bsd.uchicago.edu/data.html>). A window-based ranking process was then followed to obtain a statistic Q value for each SNP (see Supplementary Methods).

Mice. Transgenic mice expressing human T-cell receptors (Hy.2E11 and Ob.1A12) and human HLA-DR molecules were generated as described^{19,20}. We purchased 129S6/SvEvTac *Rag2*^{-/-} mice from Taconic. Clinical assessment of EAE is described in Supplementary Methods.

Adoptive transfer, apoptosis and *in vivo* proliferation assays. For apoptosis assays, splenocytes from *Hy*⁺ transgenic mice were isolated, washed with PBS and injected intravenously into recipient mice (20 × 10⁶ cells per mouse). To detect apoptotic cells, splenocytes from recipient mice were stained with antibodies against TCR Vβ4 (recognizing the β-chain of *Hy* TCR) and against CD4, together with Annexin-V and propidium iodide, and analysed by flow cytometry. For *in vivo* proliferation assays, splenocytes from *Hy*⁺ transgenic mice were labelled with 5 μM carboxyfluorescein diacetate succinimidyl ester (CFSE; Molecular Probes) in 0.1% albumin in PBS at 37°C for 10 min and washed before injection into recipient mice. To measure proliferation of transferred T cells, splenocytes from the recipient mice were stained with antibodies against CD4 and CFSE dilution was analysed by flow cytometry.

Statistical analysis of murine data. Statistical differences in EAE incidence, onset and severity, and in frequencies of *Hy*⁺CD4⁺ T cells were determined by a χ²-test or Fisher's exact test, and Student's *t*-test or the Mann-Whitney rank sum test, respectively. Values of *P* ≤ 0.05 were considered statistically significant.

Received 26 June; accepted 4 August 2006.

Published online 27 September 2006.

- Lincoln, M. R. *et al.* A predominant role for the HLA class II region in the association of the MHC region with multiple sclerosis. *Nature Genet.* **37**, 1108–1112 (2005).
- Oksenberg, J. R. *et al.* Mapping multiple sclerosis susceptibility to the HLA-DR locus in African Americans. *Am. J. Hum. Genet.* **74**, 160–167 (2004).
- Sospedra, M. & Martin, R. Immunology of multiple sclerosis. *Annu. Rev. Immunol.* **23**, 683–747 (2005).
- Fernandez-Vina, M. A. *et al.* Alleles at four HLA class II loci determined by oligonucleotide hybridization and their associations in five ethnic groups. *Immunogenetics* **34**, 299–312 (1991).
- Gao, X. J. & Serjeantson, S. W. Heterogeneity in HLA-DR2-related DR, DQ haplotypes in eight populations of Asia-Oceania. *Immunogenetics* **34**, 401–408 (1991).
- Mehra, N. K. *et al.* Analysis of HLA-DR2-associated polymorphisms by oligonucleotide hybridization in an Asian Indian population. *Hum. Immunol.* **32**, 246–253 (1991).
- Moraes, M. E., Fernandez-Vina, M. & Stastny, P. DNA typing for class II HLA antigens with allele-specific or group-specific amplification. IV. Typing for alleles of the HLA-DR2 group. *Hum. Immunol.* **31**, 139–144 (1991).
- Gao, X. J. *et al.* DNA typing for HLA-DR, and -DP alleles in a Chinese population using the polymerase chain reaction (PCR) and oligonucleotide probes. *Tissue Antigens* **38**, 24–30 (1991).
- Lee, A. *et al.* Heterogeneity of HLA-DR2 haplotypes in Caucoid Americans, African Americans, Chinese Americans, Native Americans and Xiamen Chinese. *Eur. J. Immunogenet.* **26**, 275–280 (1999).
- Song, E. Y., Kang, S. J., Lee, Y. J. & Park, M. H. HLA-DR2-associated DRB1 and DRB5 alleles and haplotypes in Koreans. *Hum. Immunol.* **61**, 937–941 (2000).
- Miretti, M. M. *et al.* A high-resolution linkage-disequilibrium map of the human major histocompatibility complex and first generation of tag single-nucleotide polymorphisms. *Am. J. Hum. Genet.* **76**, 634–646 (2005).
- Bateson, W. *Mendel's Principles of Heredity* Ch. IV, 79 (Cambridge Univ. Press, Cambridge, 1909).
- Phillips, P. C. The language of gene interaction. *Genetics* **149**, 1167–1171 (1998).
- Walsh, E. C. *et al.* An integrated haplotype map of the human major histocompatibility complex. *Am. J. Hum. Genet.* **73**, 580–590 (2003).
- Altshuler, D. *et al.* A haplotype map of the human genome. *Nature* **437**, 1299–1320 (2005).
- Fogdell, A., Hillert, J., Sachs, C. & Olerup, O. The multiple sclerosis- and

- narcolepsy-associated HLA class II haplotype includes the *DRB5*0701* allele. *Tissue Antigens* **46**, 333–336 (1995).
17. Wucherpfennig, K. W. & Strominger, J. L. Molecular mimicry in T cell-mediated autoimmunity: viral peptides activate human T cell clones specific for myelin basic protein. *Cell* **80**, 695–705 (1995).
 18. Ota, K. *et al.* T-cell recognition of an immunodominant myelin basic protein epitope in multiple sclerosis. *Nature* **346**, 183–187 (1990).
 19. Lang, H. L. *et al.* A functional and structural basis for TCR cross-reactivity in multiple sclerosis. *Nature Immunol.* **3**, 940–943 (2002).
 20. Madsen, L. S. *et al.* A humanized model for multiple sclerosis using HLA-DR2 and a human T-cell receptor. *Nature Genet.* **23**, 343–347 (1999).
 21. Green, D. R., Droin, N. & Pinkoski, M. Activation-induced cell death in T cells. *Immunol. Rev.* **193**, 70–81 (2003).
 22. Suvannavejh, G. C., Dal Canto, M. C., Matis, L. A. & Miller, S. D. Fas-mediated apoptosis in clinical remissions of relapsing experimental autoimmune encephalomyelitis. *J. Clin. Invest.* **105**, 223–231 (2000).
 23. Ganem, D. & Prince, A. M. Hepatitis B virus infection—natural history and clinical consequences. *N. Engl. J. Med.* **350**, 1118–1129 (2004).
 24. Ross, A. G. *et al.* Schistosomiasis. *N. Engl. J. Med.* **346**, 1212–1220 (2002).
 25. Zinkernagel, R. M., Pfau, C. J., Hengartner, H. & Althage, A. Susceptibility to murine lymphocytic choriomeningitis maps to class I MHC genes—a model for MHC/disease associations. *Nature* **316**, 814–817 (1985).
 26. Sabeti, P. C. *et al.* Detecting recent positive selection in the human genome from haplotype structure. *Nature* **419**, 832–837 (2002).
 27. Voight, B. F. *et al.* A map of recent positive selection in the human genome. *PLoS Biol.* **4**, e72 (2006).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank N. Willcox for discussion. Work in the authors' laboratories is supported by the Danish and British Medical Research Councils and the European Commission Descartes Prize (L.F. and J.I.B.); the Karen Elise Jensen Foundation, the Lundbeck Foundation and the Danish Multiple Sclerosis Society (L.F.); and the Wellcome Trust, the NIH and the European Union (L.R.C. and X.K.).

Author Contributions J.W.G. and K.R.K. contributed equally to this work. J.I.B. and L.F. contributed equally to this work.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to L.F. (lars.fugger@imm.ox.ac.uk).

LETTERS

Genomic analysis of increased host immune and cell death responses induced by 1918 influenza virus

John C. Kash¹, Terrence M. Tumpey², Sean C. Proll³, Victoria Carter³, Olivia Perwitasari¹, Matthew J. Thomas³, Christopher F. Basler⁴, Peter Palese⁴, Jeffery K. Taubenberger^{5†}, Adolfo García-Sastre⁴, David E. Swayne⁶ & Michael G. Katze^{1,3}

The influenza pandemic of 1918–19 was responsible for about 50 million deaths worldwide¹. Modern histopathological analysis of autopsy samples from human influenza cases from 1918 revealed significant damage to the lungs with acute, focal bronchitis and alveolitis associated with massive pulmonary oedema, haemorrhage and rapid destruction of the respiratory epithelium². The contribution of the host immune response leading to this severe pathology remains largely unknown. Here we show, in a comprehensive analysis of the global host response induced by the 1918 influenza virus, that mice infected with the reconstructed 1918 influenza virus displayed an increased and accelerated activation of host immune response genes associated with severe pulmonary pathology. We found that mice infected with a virus containing all eight genes from the pandemic virus showed marked activation of pro-inflammatory and cell-death pathways by 24 h after infection that remained unabated until death on day 5. This was in contrast with smaller host immune responses as measured at the genomic level, accompanied by less severe disease pathology and delays in death in mice infected with influenza viruses containing only subsets of 1918 genes. The results indicate a cooperative interaction between the 1918 influenza genes and show that study of the virulence of the 1918 influenza virus requires the use of the fully reconstructed virus. With recent concerns about the introduction of highly pathogenic avian influenza viruses into humans and their potential to cause a worldwide pandemic with disastrous health and economic consequences, a comprehensive understanding of the global host response to the 1918 virus is crucial. Moreover, understanding the contribution of host immune responses to virulent influenza virus infections is an important starting point for the identification of prognostic indicators and the development of novel antiviral therapies.

To study the virulence and pathogenesis of the 1918 influenza virus, mice were infected intranasally with a contemporary human influenza A/Texas/36/91 H1N1 virus (Tx91), Tx91 virus containing the 1918 HA and NA genes (2:6 1918), the 1918 HA, NA, M, NP and NS genes (5:3 1918), or the reconstructed 1918 (H1N1) recombinant virus (r1918). Mice infected with r1918 showed the most severe weight loss and the earliest mean time of death of all viruses tested (Supplementary Information). Moreover, titration of virus recovered from lung tissue showed that mice infected with r1918 had about tenfold higher levels of virus on days 1 and 3 after infection than those infected with 2:6 1918 or 5:3 1918. By day 5 after infection there was a slight decrease in the lung titres of r1918-infected mice, and mice infected with 2:6 1918, 5:3 1918 and r1918 showed similar lung

titres of about 10^7 50% embryo-infective doses per ml. Our results were consistent with those in previous reports³: most mice infected with r1918 were dead by day 5 after infection.

Previous studies showed increased lung pathology in mice infected with influenza viruses bearing genes from the 1918 virus^{3–5}. A more detailed histological analysis of mice infected with 5:3 1918, 2:6 1918 or r1918 viruses revealed differences in pathology during the course of infection. Whereas mice infected with Tx91 reassortant virus had minimal to mild diffuse histiocytic alveolitis without neutrophils or

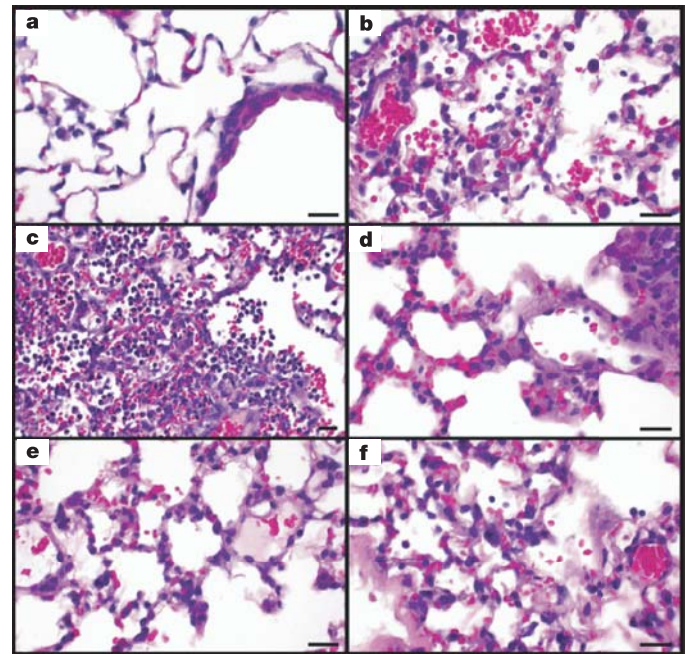


Figure 1 | Photomicrographs of lung sections from mice infected with influenza A virus, stained with haematoxylin and eosin. **a**, Mild histiocytic alveolitis with normal bronchiole, Tx/91 virus, day 3. **b**, Moderate diffuse alveolitis, predominantly neutrophilic, but with macrophages and associated haemorrhage and oedema, r1918 virus, day 3. **c**, Severe focal alveolitis with neutrophils distending affected alveoli, r1918 virus, day 3. **d, e**, Moderate histiocytic alveolitis and some neutrophils and oedema, 5:3 1918 (**d**) and 2:6 1918 (**e**) viruses, day 3. **f**, Severe diffuse histiocytic alveolitis with severe oedema, r1918 virus, day 5. Scale bar, 20 μ m.

¹Department of Microbiology, University of Washington School of Medicine, Seattle, Washington 98195, USA. ²Influenza Branch, DVRD, NCID, Centers for Disease Control and Prevention, Atlanta, Georgia 30333, USA. ³Washington National Primate Research Center, Seattle, Washington 98195, USA. ⁴Department of Microbiology, Mount Sinai School of Medicine, New York, New York 10029, USA. ⁵Department of Molecular Pathology, Department of Cellular Pathology and Genetics, Armed Forces Institute of Pathology, Rockville, Maryland 20850, USA. ⁶Southeast Poultry Research Laboratory, Agricultural Research Laboratory, US Department of Agriculture, Athens, Georgia 30606, USA.

[†]Present address: Laboratory of Infectious Diseases, NIAID, NIH, Bethesda, Maryland 20892, USA.

lesions in the bronchi or bronchioles (Fig. 1a), mice infected with 5:3 1918, 2:6 1918 or r1918 viruses had pulmonary lesions that consisted of mild to severe necrotizing bronchitis with accompanying moderate to severe alveolitis and associated moderate to severe alveolar oedema (Fig. 1b–f). However, the character of the alveolitis varied between the 1918 viruses and the sampling periods. At day 3, the r1918-infected mice had acute moderate to severe diffuse alveolitis with a prominent neutrophilic component and severe alveolar oedema (Fig. 1b). For some lungs, alveolitis was multifocal with intense neutrophilic infiltration that distended the alveoli (Fig. 1c). By contrast, mice infected with 5:3 1918 and 2:6 1918 viruses had moderate to severe alveolitis, but histiocytes predominated over neutrophils and the alveolar oedema was only moderately severe (Fig. 1d, e). At day 5 the pulmonary response in r1918-infected mice had changed to subacute and was more severe with diffuse alveolitis retaining the severe oedema, but changing to a predominance of histiocytes with fewer neutrophils (Fig. 1f). The bronchial and bronchiolar lesions were similar to those at day 3. However, in some r1918-infected lungs, an acute inflammatory response continued, as was evident from multiple scattered foci of intense neutrophilic alveolitis. In contrast, some mice infected with 5:3 1918 and 2:6 1918 viruses showed slightly less alveolitis, alveolar oedema, bronchitis and peribronchitis on day 5 than on day 3. These results show that the pathology induced by r1918 was distinct from that observed in mice infected with 5:3 1918 and 2:6 1918 viruses, indicating that cooperation between the individual 1918 viral proteins might be required for the induction of maximum pathology.

To understand why 1918 influenza virus showed extreme virulence in mice, we used gene expression analysis to provide a global view of the host response in lungs of infected mice. Expression microarray analysis was performed by taking equal masses of total RNA isolated from lungs of mice infected with strain Tx91, 2:5 1918, 5:3 1918 or r1918 and comparing them with a common reference sample prepared from lungs of mock-infected mice. Lungs from three mice were harvested for each virus on days 1, 3 and 5 after infection (total $n = 36$). Initially, our analysis was directed to genes associated with the activation of key immune cells known to be important for the response to influenza virus infection, including T helper 1 (T_H1), natural killer, macrophage and neutrophil cells (shown in Fig. 2). This analysis revealed a significant increase in the expression of genes associated with these immune cell populations in lungs of mice infected with r1918. Expression patterns of cytokines, chemokines and apoptosis-related genes showed the most significant and earliest activation during r1918 infection (Supplementary Information). The increased expression of immune-cell-related genes was correlated with virulence, because strain r1918 induced the most significant and earliest expression of immune-response-related genes, followed in magnitude and occurrence by 2:6 1918, 5:3 1918 and Tx91. Because many innate response-related genes showed saturation of the fluorescence signal by expression microarray, we measured the messenger RNA and protein expression of several key cytokines, including tumour necrosis factor (TNF), by quantitative real-time polymerase chain reaction (PCR) and enzyme-linked immunosorbent assay (ELISA). This analysis demonstrated that r1918 infections consistently resulted in a higher and prolonged expression of these mRNAs and proteins (Supplementary Information). It was shown recently⁶ that increased production of cytokines and chemokines during infection with 2:6 1918 could be attributed to the presence of neutrophils and alveolar macrophages in the lungs of infected mice. Our studies showed that r1918 infections also resulted in increased neutrophil and macrophage recruitment, accompanied by a significant increase in the expression of cytokine and chemokine genes. Expression of genes related to T helper 2 (T_H2) cell and nitric oxide metabolism showed similar patterns in all viral infections on days 1, 3 and 5, indicating that these responses might not have been primary mediators of r1918 virulence (Supplementary Information). The cellular origins of the genes identified by microarray are

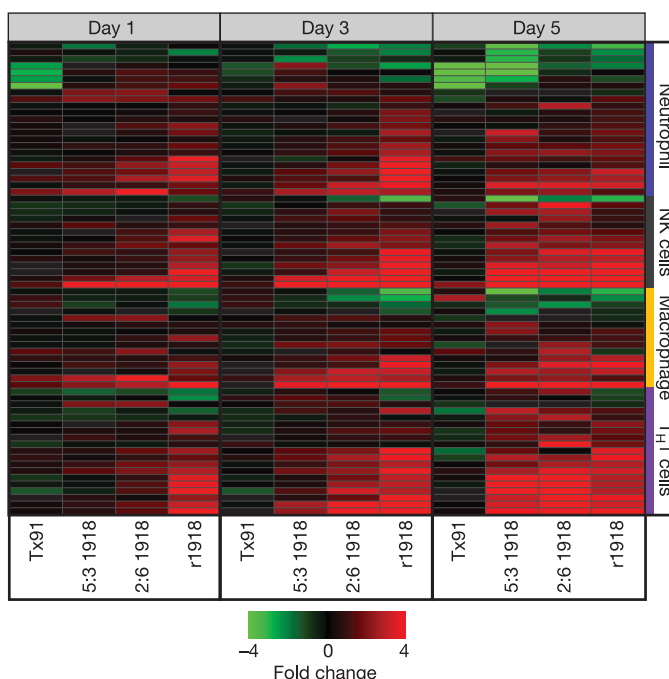


Figure 2 | Increased and earlier expression of genes associated with the activation of key immune cells in mouse lung infected with r1918 influenza virus. Expression of selected immune-cell-related genes in the lungs of mice infected with Tx91, 2:6 1918, 5:3 1918 or r1918 viruses on days 1, 3 and 5 after infection, as determined by expression microarray analysis. For each infection point, the data presented are the error-weighted average expression changes calculated from four technical replicate arrays performed on three individual mice ($n = 12$ total). Genes shown in red were upregulated and genes shown in green were downregulated in infected relative to mock-infected mouse lung. NK, natural killer; T_H1 , T helper 1.

unknown, but taken together these data show that r1918 produced the most severe and rapid disease in mice, which could be correlated with an increased expression of inflammatory response genes similar to that reported for highly pathogenic avian influenza viruses^{7,8}.

To characterize the functional consequences of gene expression changes associated with infection with r1918, we performed pathway analysis of the gene expression data with Ingenuity Pathways Analysis (see Methods). As shown in Fig. 3, pathway analysis of the expression array data revealed that infection with r1918 virus resulted in the most significant activation of death receptor, interleukin-6, type I interferon and Toll-like receptor responses. These data show that the aggregate host response to r1918 virus infection results in an increase in pro-inflammatory and cell death responses. We observed an inverse correlation between influenza virulence and G2/M checkpoint control and glutathione metabolism. The increased expression of cell-cycle and glutathione metabolism genes in Tx91 infections could be important components in the successful clearance of influenza virus infection and tissue remodelling without severe immunopathology. We also suggest that activation of these responses reflects the importance of protecting cells from damage by oxygen radicals and the need for the proliferation and repopulation of damaged or killed cells in the respiratory epithelium.

To explore the host responses associated with the severe immunopathology of infections with r1918 virus, we characterized the functional relationship between genes that were induced more than twofold ($P < 0.01$) only in r1918 virus infections. As shown in Fig. 4, many of these genes that were uniquely induced more than twofold in r1918-infected mouse lung compared with controls were involved in cell death and inflammatory responses and have been reported to have direct and/or indirect interactions. Infections with r1918 resulted in a statistically significant coordinated activation of

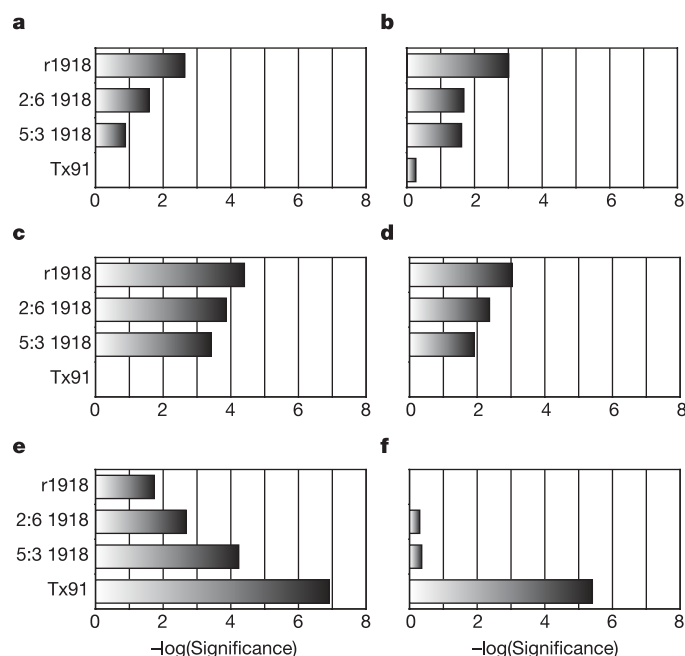


Figure 3 | Increased activation of inflammatory and death receptor gene expression by r1918 influenza virus. **a–d**, Analysis of the biological response pathways that showed most significant activation during r1918 infection, including death receptor (**a**), interleukin-6 (**b**), interferon (**c**) and Toll-like receptor (**d**) pathways. **e, f**, The biological responses most significantly induced in the non-lethal Tx91 infections: G2/M checkpoint control (**e**) and glutathione metabolism (**f**). Pathway analysis was performed on genes that were induced more than twofold ($P < 0.01$) in any of the infections compared with uninfected controls. For this purpose, Ingenuity Pathway Analysis was used, which analyses the RNA expression data in the context of known biological response and regulatory networks, as well as other higher-order response pathways to assign functional information and biological relevance. These results are shown as the negative logarithm of significance, which is a statistical score and is a measure of the likelihood of the genes in a given network being found together as a result of chance, as determined by Fisher's exact test (see Methods for details).

multiple cell-death-related genes, particularly those related to death receptor responses. This is significant because activation of TNF-related pathways has been shown to be associated with H5N1 influenza viral infections of macrophages⁹. Moreover, r1918 infection-induced expression of mRNAs for Fas, caspase-8 and caspase-9 is significant because of the central roles of these proteins in the induction of apoptosis through the death-receptor and mitochondrial apoptosis pathways. Caspase-9 is activated by the release of cytochrome *c* from permeabilized mitochondria, and it has been reported that an influenza viral protein expressed from an alternative reading frame of the PB1 gene segment (PB1-F2) is trafficked to mitochondria and can induce apoptosis¹⁰. It has recently been reported¹¹ that PB1-F2 associates with components of the permeability transition pore complex and sensitizes cells to apoptotic stimuli, including TNF- α , resulting in mitochondrial permeabilization and the release of cytochrome *c*. We believe our data showing increased cell death responses in r1918-infected mice may be related to the effects of the 1918 PB1-F2 protein. However, these observations do not rule out additional functions of the 1918 polymerase proteins, because these genes are also important virulence factors for highly pathogenic avian influenza viruses^{12–17}.

Taken together, our results indicate that enhanced inflammatory and cell death responses might be contributors to severe immunopathology associated with r1918 influenza virus infections, although it is not possible to completely exclude the possibility that increased immune and cell death responses may be a consequence rather than the cause of enhanced pathology of r1918 infection. Our study also

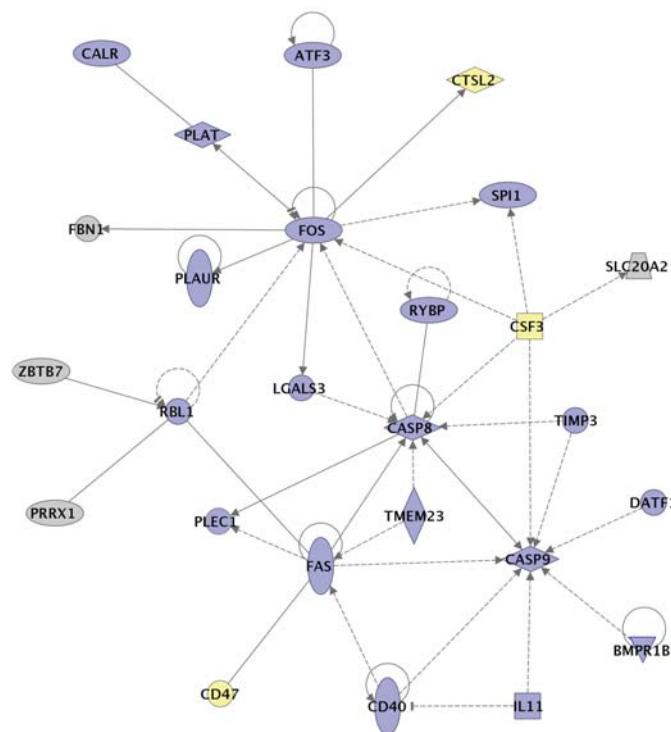


Figure 4 | Functional relationships of activated cell death responses during infection with r1918 influenza virus. Biological network of selected genes that were only induced more than twofold ($P < 0.01$) in the r1918-infected mouse lung compared with uninfected controls. This diagram shows the direct (solid lines) and indirect (dashed lines) interactions reported for these cell-death-related (blue shading) and immune-response-related (yellow shading) genes; grey denotes genes with multiple and/or undefined biological functions. Biological network analysis was performed by using Ingenuity Pathway Analysis and showed statistical significance with Fisher's exact test to calculate a P -value determining the probability that each biological function and/or disease assigned to that data set was due to chance alone (see Methods for details).

shows the synergistic effect of the 1918 gene segments and that exceptional virulence and exaggerated host global gene expression response may be dependent on the complete constellation of 1918 genes. We acknowledge that a more complete understanding of r1918 virulence will require experiments in additional animal models, including ferrets, pigs and non-human primates. Characterization of the global host response to other highly virulent viruses will be important in understanding the role of the host in influenza virus pathology, and we are currently comparing these results with findings from similar studies of animals infected with highly pathogenic avian influenza viruses.

METHODS

Recombinant influenza viruses. Recombinant viruses were generated with a reverse genetics system¹⁸. The low cell (C1 or C2) passage influenza A H1N1 virus stocks were grown in Madin–Darby canine kidney cells. The reconstructed strain of human influenza H1N1 A/Texas/36/91 (Tx/91) virus has been characterized previously³ and causes common influenza symptoms in humans in experimental settings¹⁹. The r1918 influenza virus genes correspond to the published sequence of influenza A/South Carolina/1/18 (H1N1) virus HA open reading frame²⁰, and influenza A/Brevig Mission/1/18 (H1N1) virus NA, M, NS, NP, PA, PB1 and PB2 open reading frames^{12,21–24}. The 1918 recombinant viruses contain the 5' and 3' non-coding regions corresponding to influenza A/WSN/33 (H1N1) virus.

Mouse experiments. Female BALB/c mice (Harlan Laboratories) 6–7 weeks old were anesthetized with ketamine–xylazine (1.98 and 0.198 mg per mouse, respectively) and inoculated intranasally with 10^6 plaque-forming units of the indicated virus. Because of the virulence of r1918 in mice, increased numbers of

animals were infected with r1918 to allow the collection of samples at day 5 after infection³. Lungs were harvested at days 1, 3 and 5 after infection for viral titration and histopathology, and extraction of total RNA was performed from a 10% (w/v) lung homogenate as described previously⁵. All animal work was performed in a specially separated negative-pressure HEPA (high-efficiency particulate air)-filtered biosafety level laboratory (BSL3) with enhancements. All personnel wore half-body Racal hoods with backpack HEPA-filtered air supplies⁵.

Histopathology. Three mice from each group were killed at days 3 and 5 after infection. A 5-mm lung piece from the ventral end of each left lobe was collected for histopathology. Tissues were fixed with standard procedures in 10% neutral-buffered formalin solution, sectioned, and stained with haematoxylin and eosin.

Expression microarray analysis and bioinformatics. Total RNA isolation and mRNA amplification were performed on equal masses of total RNA isolated from lungs of infected mice as described previously²⁵ and were compared with an equal-mass pool of RNA from lungs of ten individual mock-infected mice. For each infection group, expression oligonucleotide arrays ($n = 4$ per gene) were performed on RNA isolated from lung tissue from three individual animals ($n = 12$ total). The data presented are the error-weighted average changes in expression calculated from four technical replicate arrays performed on three individual mice ($n = 12$ final for each gene). Probe labelling and microarray slide hybridization were performed with mouse oligonucleotide arrays (G4121A; Agilent Technologies)²⁵. All data were entered into a custom-designed Oracle 9i backed relational database, Expression Array Manager, and were then uploaded into Rosetta Resolver System 5.0 (Rosetta Biosoftware) and Spotfire Decision Site 8.1 (Spotfire). All primary expression microarray data are available at the University of Washington's Public Microarray Data Download site (<http://expression.microslu.washington.edu>), in accordance with the proposed MIAME (minimum information about a microarray experiment) standards²⁶.

Functional analysis of statistically significant gene expression changes was performed with Ingenuity Pathways Analysis (Ingenuity Systems). This software analyses RNA expression data in the context of known biological response and regulatory networks as well as other higher-order response pathways. Ingenuity functional analysis identified biological functions and/or diseases that were most significant. Genes from the data set that met the twofold ($P < 0.01$) change cutoff and were associated with biological functions in the Ingenuity Pathways Knowledge Base were considered for analysis. For all analyses, Fisher's exact test was used to calculate a P -value determining the probability that each biological function assigned to that data set was due to chance alone. In the functional network shown in Fig. 4, genes are represented as nodes, and the biological relationship between two nodes is represented as an edge (line). All edges are supported by at least one published reference or from canonical information stored in the Ingenuity Pathways Knowledge Base. For these analyses, Fisher's exact test was used to calculate a P -value determining the probability that each biological function and/or disease assigned to that data set was due to chance alone.

Real-time PCR assays. Quantitative real-time PCR was used to validate changes in gene expression²⁵. Primer and probe sets for each of the target sequences were chosen from the Applied Biosystems Assays-on-Demand product list. Quantification of each gene, relative to the calibrator, was performed with the equation $2^{-\Delta\Delta C_t}$ within the Applied Biosystems Sequence Detection Software version 1.2.2.

Cytokine and chemokine quantification. To determine levels of cytokine or chemokine proteins in lung, at each time point examined three mice from each group were exsanguinated from the axilla and killed; lung tissues were then removed from naive and infected mice (days 1, 3 and 5 after infection). Individual whole lung samples were immediately frozen at -70°C . At the time of analysis, tissues were thawed, homogenized in 1 ml of cold PBS and centrifuged at 150g for 5 min. With the use of ELISA kits from R&D Systems, the clarified cell lysates were assayed for mouse IP-10 (CXCL10), TNF- α , macrophage inflammatory protein (MIP)-1 α (CCL3) and MIP-2 (CXCL2).

Received 19 June; accepted 18 August 2006.

Published online 27 September 2006.

- Johnson, N. P. & Mueller, J. Updating the accounts: global mortality of the 1918–1920 'Spanish' influenza pandemic. *Bull. Hist. Med.* **76**, 105–115 (2002).
- Taubenberger, J. K. & Morens, D. M. 1918 influenza: the mother of all pandemics. *Emerg. Infect. Dis.* **12**, 15–22 (2006).
- Tumpey, T. M. *et al.* Characterization of the reconstructed 1918 Spanish influenza pandemic virus. *Science* **310**, 77–80 (2005).
- Kobasa, D. *et al.* Enhanced virulence of influenza A viruses with the haemagglutinin of the 1918 pandemic virus. *Nature* **431**, 703–707 (2004).

- Kash, J. C. *et al.* Global host immune response: pathogenesis and transcriptional profiling of type A influenza viruses expressing the hemagglutinin and neuraminidase genes from the 1918 pandemic virus. *J. Virol.* **78**, 9499–9511 (2004).
- Tumpey, T. M. *et al.* Pathogenicity of influenza viruses with genes from the 1918 pandemic virus: functional roles of alveolar macrophages and neutrophils in limiting virus replication and mortality in mice. *J. Virol.* **79**, 14933–14944 (2005).
- Zhou, J. *et al.* Differential expression of chemokines and their receptors in adult and neonatal macrophages infected with human or avian influenza viruses. *J. Infect. Dis.* **194**, 61–70 (2006).
- Cheung, C. Y. *et al.* Induction of proinflammatory cytokines in human macrophages by influenza A (H5N1) viruses: a mechanism for the unusual severity of human disease? *Lancet* **360**, 1831–1837 (2002).
- Zhou, J. *et al.* Functional tumor necrosis factor-related apoptosis-inducing ligand production by avian influenza virus-infected macrophages. *J. Infect. Dis.* **193**, 945–953 (2006).
- Chen, W. *et al.* A novel influenza A virus mitochondrial protein that induces cell death. *Nature Med.* **7**, 1306–1312 (2001).
- Zamarin, D., Garcia-Sastre, A., Xiao, X., Wang, R. & Palese, P. Influenza virus PB1–F2 protein induces cell death through mitochondrial ANT3 and VDAC1. *PLoS Pathog.* **1**, e4 (2005).
- Taubenberger, J. K. *et al.* Characterization of the 1918 influenza virus polymerase genes. *Nature* **437**, 889–893 (2005).
- Salomon, R. *et al.* The polymerase complex genes contribute to the high virulence of the human H5N1 influenza virus isolate A/Vietnam/1203/04. *J. Exp. Med.* **203**, 689–697 (2006).
- Gabriel, G. *et al.* The viral polymerase mediates adaptation of an avian influenza virus to a mammalian host. *Proc. Natl Acad. Sci. USA* **102**, 18590–18595 (2005).
- Shinya, K. *et al.* PB2 amino acid at position 627 affects replicative efficiency, but not cell tropism, of Hong Kong H5N1 influenza A viruses in mice. *Virology* **320**, 258–266 (2004).
- Hatta, M., Gao, P., Halfmann, P. & Kawaoka, Y. Molecular basis for high virulence of Hong Kong H5N1 influenza A viruses. *Science* **293**, 1840–1842 (2001).
- Katz, J. M. *et al.* Molecular correlates of influenza A H5N1 virus pathogenesis in mice. *J. Virol.* **74**, 10807–10810 (2000).
- Fodor, E. *et al.* Rescue of influenza A virus from recombinant DNA. *J. Virol.* **73**, 9679–9682 (1999).
- Hayden, F. G. *et al.* Local and systemic cytokine responses during experimental human influenza A virus infection. Relation to symptom formation and host defense. *J. Clin. Invest.* **101**, 643–649 (1998).
- Reid, A. H., Fanning, T. G., Hultin, J. V. & Taubenberger, J. K. Origin and evolution of the 1918 'Spanish' influenza virus hemagglutinin gene. *Proc. Natl Acad. Sci. USA* **96**, 1651–1656 (1999).
- Reid, A. H., Fanning, T. G., Janczewski, T. A., McCall, S. & Taubenberger, J. K. Characterization of the 1918 'Spanish' influenza virus matrix gene segment. *J. Virol.* **76**, 10717–10723 (2002).
- Reid, A. H., Fanning, T. G., Janczewski, T. A. & Taubenberger, J. K. Characterization of the 1918 'Spanish' influenza virus neuraminidase gene. *Proc. Natl Acad. Sci. USA* **97**, 6785–6790 (2000).
- Basler, C. F. *et al.* Sequence of the 1918 pandemic influenza virus nonstructural gene (NS) segment and characterization of recombinant viruses bearing the 1918 NS genes. *Proc. Natl Acad. Sci. USA* **98**, 2746–2751 (2001).
- Reid, A. H., Fanning, T. G., Janczewski, T. A., Lourens, R. M. & Taubenberger, J. K. Novel origin of the 1918 pandemic influenza virus nucleoprotein gene. *J. Virol.* **78**, 12462–12470 (2004).
- Kash, J. C. *et al.* Global suppression of the host antiviral response by Ebola- and Marburgviruses: increased antagonism of the type I interferon response is associated with enhanced virulence. *J. Virol.* **80**, 3009–3020 (2006).
- Brazma, A. *et al.* Minimum information about a microarray experiment (MIAME)—toward standards for microarray data. *Nature Genet.* **29**, 365–371 (2001).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank K. A. Walters, M. J. Korth, J. Fornek and B. Paepke for discussions and assistance in manuscript preparation. This work was partly supported by grants from the National Institutes of Health (to A.G.S., P.P., J.K.T. and M.G.K.) and by the Ellison Medical Foundation (P.P. and C.F.B.).

Author Information Raw data on expression microarrays are available at <http://expression.microslu.washington.edu>. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to J.K. (jkash@u.washington.edu).

LETTERS

Sodium-dependent uptake of inorganic phosphate by the intracellular malaria parasite

Kevin J. Saliba^{1,2*}, Rowena E. Martin^{1*}, Angelika Bröer¹, Roselani I. Henry¹, C. Siobhan McCarthy¹, Megan J. Downie¹, Richard J. W. Allen¹, Kylie A. Mullin³, Geoffrey I. McFadden³, Stefan Bröer¹ & Kiaran Kirk¹

As the malaria parasite, *Plasmodium falciparum*, grows within its host erythrocyte it induces an increase in the permeability of the erythrocyte membrane to a range of low-molecular-mass solutes, including Na^+ and K^+ (ref. 1). This results in a progressive increase in the concentration of Na^+ in the erythrocyte cytosol^{2,3}. The parasite cytosol has a relatively low Na^+ concentration^{2,4} and there is therefore a large inward Na^+ gradient across the parasite plasma membrane. Here we show that the parasite exploits the Na^+ electrochemical gradient to energize the uptake of inorganic phosphate (P_i), an essential nutrient. P_i was taken up into the intracellular parasite by a Na^+ -dependent transporter, with a stoichiometry of $2\text{Na}^+:\text{P}_i$ and with an apparent preference for the monovalent over the divalent form of P_i . A P_i transporter (PfPiT) belonging to the PiT family was cloned from the parasite and localized to the parasite surface. Expression of PfPiT in *Xenopus* oocytes resulted in Na^+ -dependent P_i uptake with characteristics similar to those observed for P_i uptake in the parasite. This study provides new insight into the significance of the malaria-parasite-induced alteration of the ionic composition of its host cell.

P_i is an important nutrient in cell metabolism and is required for the synthesis of DNA, RNA and numerous phosphorylated metabolic intermediates. Removal of P_i from the growth medium resulted in a $92 \pm 3\%$ ($n = 5$; mean $\pm$ s.e.m.; $P < 0.0001$) decrease in parasite proliferation (measured over 96 h). The intraerythrocytic parasite is therefore dependent on an external supply of P_i to maintain normal growth.

The mechanism by which the parasite takes up P_i from the host erythrocyte was investigated by using parasites 'isolated' from their host cells with saponin⁵. Initial uptake experiments were performed in 'de-energized' (glucose-depleted and therefore ATP-depleted) cells to avoid complications from metabolism. The influx of $^{33}\text{P}_i$ into de-energized parasites was linear with time for more than 20 min, with the $^{33}\text{P}_i$ distribution ratio ($[^{33}\text{P}_i]_{\text{in}}/[^{33}\text{P}_i]_{\text{out}}$) increasing to more than 40. Replacement of Na^+ in the medium with either K^+ (Fig. 1a) or choline (not shown) decreased P_i influx by more than 95%, which is consistent with the involvement of a Na^+ -dependent transporter. P_i influx showed a sigmoidal dependence on the Na^+ concentration (Fig. 1b) with a Hill coefficient of 2.1 ± 0.2 ($n = 9$; mean $\pm$ s.e.m.), consistent with two Na^+ ions binding to the transporter for each P_i molecule. The stoichiometry was confirmed by a 'static head' experiment⁶ in which $^{33}\text{P}_i$ was allowed to equilibrate in the absence of a Na^+ gradient, then the suspension was diluted in $^{33}\text{P}_i$ -free medium containing a range of Na^+ concentrations. The $[\text{Na}^+]_{\text{out}}$ required to keep the intracellular $[^{33}\text{P}_i]$ at the same level as in parasites maintained in the original loading solution was $69 \pm 12 \text{ mM}$ ($n = 7$, mean $\pm$ s.e.m.), which is consistent with a

stoichiometry of about 2 (1.7 ± 0.3) Na^+ ions being transported per P_i molecule (Fig. 1c).

The apparent K_m for the uptake of P_i ($K_m(\text{P}_i)$) decreased with increasing extracellular $[\text{Na}^+]$, with values of 445 ± 31 , 258 ± 16 and $106 \pm 8 \mu\text{M}$ at Na^+ concentrations of 5, 15 and 100 mM, respectively ($P < 0.01$; Fig. 1d). Thus, as the concentration of Na^+ in the erythrocyte increases with parasite maturation^{2,3}, the efficiency of the parasite's P_i uptake mechanism increases.

The apparent $K_m(\text{P}_i)$ also showed a strong dependence on the pH of the medium, decreasing as the extracellular pH decreased (Fig. 1e). P_i has a pK_a of 6.8 under physiological conditions⁷, and as extracellular pH varies so, too, do the relative proportions of monovalent (H_2PO_4^-) and divalent (HPO_4^{2-}) P_i . If P_i influx is plotted as a function of the (calculated) concentration of H_2PO_4^- , similar apparent K_m values are obtained at each of the three extracellular pH values tested (Table 1). If the same analysis is performed for HPO_4^{2-} , the apparent K_m values diverge (Table 1). These data are consistent with the influx of P_i being by means of a transporter with a strong preference for H_2PO_4^- over HPO_4^{2-} , and an apparent $K_m(\text{P}_i)$ that varies little with pH.

A stoichiometry of $2\text{Na}^+:\text{H}_2\text{PO}_4^-$ predicts that net uptake of P_i should be electrogenic, involving an influx of positive charge. Under physiological conditions the parasite has a large, inwardly negative plasma membrane potential⁸, which should therefore influence P_i uptake. To test this, $^{33}\text{P}_i$ influx was measured in energized (ATP-replete) parasites. Depolarization of the parasite plasma membrane with carbonylcyanide-*p*-trifluoromethoxyphenylhydrazone (FCCP) or by the plasma membrane H^+ -pump inhibitor bafilomycin A_1 (ref. 8) resulted in a significant decrease in $^{33}\text{P}_i$ influx (Fig. 1f), consistent with the predicted electrogenicity of transport.

Results of additional (efflux) experiments with isolated parasites were consistent with the transporter catalysing not only the net influx and efflux of P_i but also a non-productive exchange of intracellular P_i for extracellular P_i (Supplementary Fig. 2). These transport characteristics may be accounted for by the mechanistic model presented in Supplementary Fig. 3. $^{33}\text{P}_i$ efflux by the exchange reaction was much faster than net efflux, which is consistent with net efflux being rate limited by the reorientation of the empty transporter, whereas exchange involves the faster reorientation of the $[\text{H}_2\text{PO}_4^- + 2\text{Na}^+]$ -loaded transporter. Depolarization of the parasite plasma membrane slowed net P_i efflux but not the exchange reaction, which is consistent with the presence of a negative charge on the transporter, rendering the $[\text{H}_2\text{PO}_4^- + 2\text{Na}^+]$ -loaded transporter complex electroneutral. Exchange was slowed (but not abolished) by the removal of extracellular Na^+ , whereas the same manoeuvre enhanced net efflux. The efflux of $^{33}\text{P}_i$ into Na^+ -free medium might be explained by a partial reaction of the transporter (Supplementary

¹School of Biochemistry and Molecular Biology, ²Medical School, The Australian National University, Canberra ACT 0200, Australia. ³School of Botany, University of Melbourne, Victoria 3010, Australia.

*These authors contributed equally to this work.

Fig. 3), although the involvement of additional transport pathways cannot be excluded.

The annotated *P. falciparum* genome has a single putative plasma membrane P_i transporter (PlasmoDB accession no. MAL13P1.206)^{9,10}, designated here as PfPiT. Sequence comparisons show PfPiT to be a member of a family of H^+ - or Na^+ -coupled P_i transporters (the PiT family, or solute carrier family 20), other members of which are

present in prokaryotes, plants, fungi and animals (Supplementary Fig. 4). Phylogenetic analysis reveals that PfPiT bears a closer resemblance to Na^+ -dependent transporters from fungi and animals than to the H^+ -dependent PiT transporters of bacteria and plant chloroplasts (Fig. 2a). An alignment of the hydropathy plots for PfPiT and mouse PiT-1 (Supplementary Fig. 5) shows the strong structural homology of the parasite protein to animal PiT family members and suggests the presence of 12 transmembrane domains (Fig. 2b), consistent with the proposed topology of the human homologue PiT-2 (ref. 11). PiT-2 has both the amino and carboxy termini located at the extracellular face of the membrane¹¹, and the same is likely to be true of PfPiT.

Messenger RNA encoding the *P. falciparum* protein is expressed throughout the entire blood stage of the parasite's life cycle^{9,12,13}. The level of transcript inside the cell increases markedly between 16 and 24 h after invasion⁹, during which time the parasite-induced increase in erythrocyte $[Na^+]$ is predicted to occur³. Transfection of the parasite with PfPiT complementary DNA, tagged at the C terminus with a haemagglutinin (HA) tag, enabled the localization of the protein by immunofluorescence (Fig. 2c). Fluorescence was restricted to the parasite (Fig. 2c, top three panels), and colocalized with that obtained with an antibody against the (unrelated) plasma membrane protein PfENT1 (ref. 14) (Fig. 2c, bottom three panels).

To investigate the function of PfPiT we expressed the protein in *Xenopus laevis* oocytes. Oocytes injected with PfPiT cRNA showed a fivefold to tenfold increase in $^{33}P_i$ uptake relative to non-injected controls (Fig. 3a) or to oocytes expressing PfENT1 (data not shown). The increase was comparable in magnitude to that seen in oocytes expressing the mouse homologue mPiT-1 (data not shown). Replacement of extracellular Na^+ by *N*-methyl-D-glucamine⁺ abolished PfPiT-induced P_i uptake (Fig. 3a; $P < 0.0001$), which is consistent with PfPiT being a Na^+ -dependent transporter. An analysis of uptake as a function of phosphate concentration at pH 7.4 yielded an apparent $K_m(P_i)$ of $20 \pm 2 \mu M$ and a V_{max} of $40 \pm 1 \text{ pmol h}^{-1}$ per oocyte (Fig. 3b). The former figure is close to the K_m values obtained for the human PiT-1, rat PiT-2 and mouse PiT-2 phosphate transporters when expressed in oocytes^{15,16}, although it is lower than the apparent $K_m(P_i)$ determined in the experiments with isolated parasites under similar conditions (Table 1). Discrepancies between the apparent K_m determined for a transport protein in its native cell and that measured for the same protein expressed in a heterologous system have been observed previously^{17,18}.

As was observed in isolated parasites (Fig. 1e), the rate of P_i influx into oocytes expressing PfPiT increased with decreasing pH (Fig. 3c), which is consistent with PfPiT having a strong preference for $H_2PO_4^-$. Depolarization of the oocyte by the replacement of 50 mM NaCl with 50 mM KCl in the medium¹⁹ halved P_i uptake (Fig. 3d; $P < 0.0001$), a similar decrease to that seen for the influx of P_i into isolated parasites after membrane depolarization (Fig. 1f).

Thus, PfPiT, the only putative plasma membrane P_i transporter

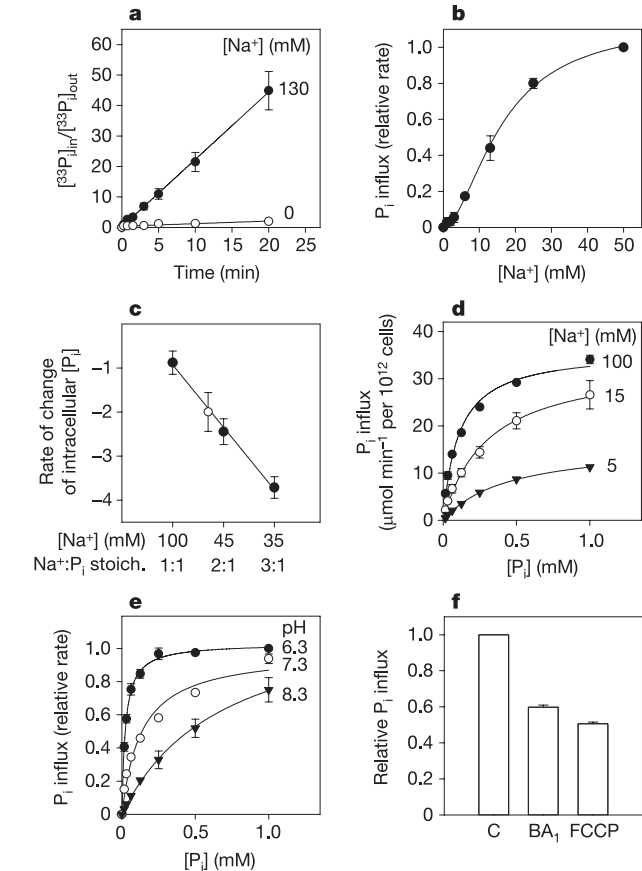


Figure 1 | P_i transport in malaria parasites 'isolated' from their host erythrocytes by permeabilization with saponin. **a**, Time courses for the uptake of $^{33}P_i$ into isolated parasites, de-energized by preincubation in the absence of glucose and suspended in 130 mM Na^+ -containing medium (filled circles) or in medium in which Na^+ was replaced iso-osmotically with K^+ (open circles). **b**, $[Na^+]$ dependence of the influx of $^{33}P_i$ into de-energized isolated parasites. **c**, Estimate of $Na^+:P_i$ stoichiometry in a 'static head' experiment in which de-energized parasites were pre-equilibrated with $^{33}P_i$ in the absence of a Na^+ gradient ($[Na^+]_{in} \approx [Na^+]_{out} = 20 \text{ mM}$), then either maintained in the loading solution (under which conditions $[^{33}P_i]_{in}$ slowly decreased) or diluted fivefold in medium of three differing Na^+ concentrations, corresponding to the $[Na^+]$ values required by a $1Na^+:1P_i$, $2Na^+:1P_i$, and $3Na^+:1P_i$ transporter to maintain the rate of decrease of $[^{33}P_i]_{in}$ the same as that in cells maintained in the original loading solution. The filled symbols show the rate of decrease of $[^{33}P_i]_{in}$ in the three solutions with differing $[Na^+]$ and the open symbol shows the rate of decrease observed in cells maintained in the original loading solution. **d**, $[P_i]$ dependence of the influx of $^{33}P_i$ into de-energized isolated parasites suspended in medium containing 100 mM (filled circles), 15 mM (open circles) or 5 mM (filled triangles) Na^+ (pH 7.4). **e**, $[P_i]$ dependence of the influx of $^{33}P_i$ into de-energized isolated parasites suspended in medium at pH 6.3 (filled circles), 7.3 (open circles) or 8.3 (filled triangles) ($[Na^+]_{out} = 130 \text{ mM}$). The corresponding kinetic constants are given in Table 1. **f**, Effect of depolarization of the parasite membrane potential by either bafilomycin A_1 (BA_1 ; 100 nM) or FCCP (10 μM) on the influx of P_i into isolated parasites. The intracellular and extracellular pH values were both 7.35. Influx rates are normalized relative to that measured in energized parasites under control (C) conditions. Results are mean $\pm$ s.e.m. from multiple ($n \geq 3$) experiments.

Table 1 | pH dependence of the kinetic parameters for P_i influx into isolated and de-energized *P. falciparum* parasites

pH	V_{max} ($\mu mol min^{-1}$ per 10^{12} parasites)	Apparent K_m (μM)		
		$[P_i]$	$[H_2PO_4^-]$	$[HPO_4^{2-}]$
6.3	27 ± 6	$30 \pm 6^*$	23 ± 4	$7 \pm 1^{**}$
7.3	22 ± 5	$118 \pm 15^*$	28 ± 4	$90 \pm 11^{**}$
8.3	28 ± 6	$621 \pm 103^*$	19 ± 3	$601 \pm 100^{**}$

The three different apparent K_m values at each pH are those estimated when the analysis was performed using the total P_i concentration ($[P_i]$, as in Fig. 1e), the calculated concentration of monovalent $H_2PO_4^-$, and the calculated concentration of divalent HPO_4^{2-} , in the extracellular medium. All values are mean $\pm$ s.e.m. for four different experiments. Asterisk, values obtained at the three different pH values differ significantly from one another (ANOVA; $P < 0.0002$). Two asterisks, values obtained at the three different pH values differ significantly from one another (ANOVA; $P < 0.0001$). No asterisk, values obtained at the three different pH values do not differ significantly from one another (ANOVA; $P = 0.268$).

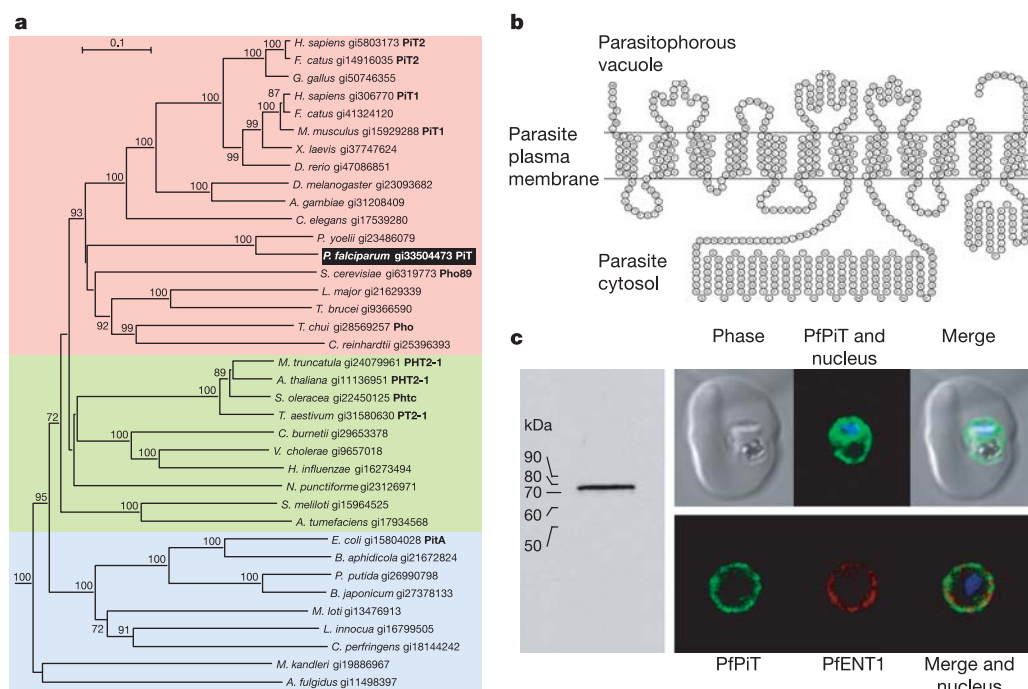


Figure 2 | The *P. falciparum* P_i transporter, PfPiT. a, Phylogenetic tree showing relationships between PfPiT and homologues in other organisms. The main branches of the tree are indicated by backgrounds as follows: red, proteins from animals, fungi and protists; green, proteins from plants and bacteria; blue, a second cluster of bacterial proteins and two Archaeal proteins. Bootstrap scores of more than 70% are shown on the branches. The scale bar represents the number of substitutions per site for a unit branch length. **b**, Topographical model of PfPiT derived from hydropathy plots (Supplementary Fig. 5), from a protein sequence alignment of the PiT family (Supplementary Fig. 4) and from topological data for the related human PiT-2 transporter¹¹. **c**, Western blot and immunolocalization of PfPiT in the

P. falciparum-infected erythrocyte. Left: western blot analysis of cells expressing PfPiT-HA. A single band corresponding to the predicted size of PfPiT-HA (about 78 kDa) was detected. Protein from erythrocytes infected with wild-type 3D7 *P. falciparum* did not bind the HA antibody (data not shown). Top row of three panels: a *P. falciparum*-infected erythrocyte expressing PfPiT-HA (green) and nucleus stained with Hoechst 33258 (blue). Within the infected cell, PfPiT-HA localizes to the parasite. Lower row of three panels: PfPiT localizes to the parasite surface, colocalizing with PfENT1 (red), a transport protein shown previously to reside on the plasma membrane of the parasite¹⁴.

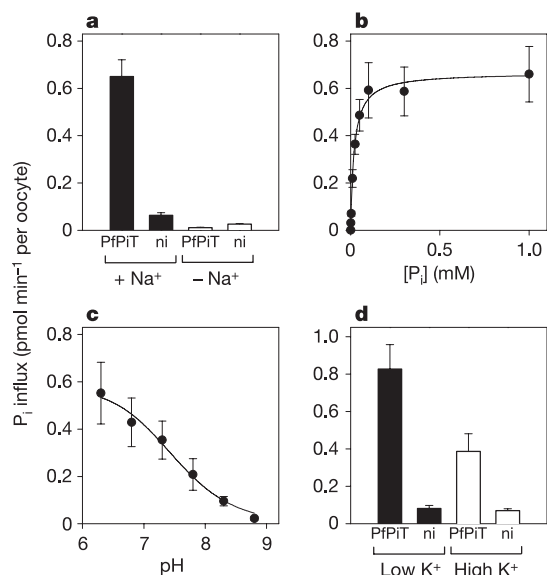


Figure 3 | Transport characteristics of PfPiT in *Xenopus* oocytes. a, Oocytes expressing PfPiT show a marked increase in $^{33}P_i$ uptake relative to non-injected (ni) oocytes. Replacement of NaCl by *N*-methyl-D-glucamine chloride in the incubation medium abolished PfPiT-induced $^{33}P_i$ uptake ($[P_i] = 10 \mu M$). **b**, $[P_i]$ dependence of PfPiT-induced P_i uptake. PfPiT-induced uptake was calculated by subtracting the uptake measured in non-injected oocytes from that in oocytes expressing PfPiT. **c**, pH dependence of PfPiT-induced P_i uptake ($[P_i] = 25 \mu M$). **d**, Depolarization of oocytes by the replacement of 50 mM NaCl with 50 mM KCl in the incubation medium¹⁹ halved P_i uptake. P_i uptake is shown as mean $\pm$ s.d. measured in seven to ten oocytes.

identified so far in the *P. falciparum* genome^{9,10}, is localized to the parasite surface, and when expressed in *X. laevis* oocytes has P_i transport characteristics similar to those of isolated parasites. The data are consistent with PfPiT having an important role in the uptake of P_i by the intracellular parasite, although the involvement of additional (Na^+ -dependent) transporters cannot be ruled out.

As an obligate intracellular parasite, *P. falciparum* inhabits an initially low- Na^+ environment, and in the original annotation of the *P. falciparum* genome it was assumed that the parasite would use H^+ -coupled rather than Na^+ -coupled transporters for the uptake of nutrients, including P_i (ref. 10). The results of the present study show that this is not necessarily the case, revealing the unexpected scenario of a Na^+ -coupled nutrient transport mechanism in an intracellular parasite. Na^+ -coupled transporters such as PfPiT provide the means for the malaria parasite to exploit the increased $[Na^+]$ in the infected erythrocyte cytosol (together with the parasite's plasma membrane potential⁸), to energize the uptake of solutes. The results of the present study (summarized in Supplementary Fig. 1) therefore provide new insight into the physiological role for the long-recognized perturbation by the parasite of the ionic composition of its host cell.

METHODS

P_i transport in isolated malaria parasites. *P. falciparum* parasites (strain FAF-6) were cultured and synchronized as described elsewhere⁵. The transport of P_i across the plasma membrane of mature 'trophozoite-stage' parasites (about 36 h after invasion) was measured in parasites 'isolated' from their host erythrocytes by permeabilization of the erythrocyte and parasitophorous vacuole membranes with saponin⁵ (Supplementary Methods). Parasite growth was monitored by measuring the incorporation of $[^3H]$ hypoxanthine over 96 h by parasites

suspended in a conventional medium (containing 5.6 mM P_i) or in an equivalent medium from which P_i was omitted.

Expression of PfPiT, mPiT-1 and PfENT1 in *X. laevis* oocytes. Genes encoding the parasite proteins PfPiT and PfENT1 and the mouse PiT-1 (mPiT-1) protein were cloned into oocyte expression vector pGem-He-Juel²⁰ (Supplementary Methods). *In vitro* transcription, oocyte preparation and cRNA injection (30 ng per oocyte) were performed as described elsewhere²¹. ³³P_i uptake was measured (4–6 days after injection) in medium that, unless specified otherwise, contained 96 mM NaCl, 2 mM KCl, 1 mM MgCl₂, 1.8 mM CaCl₂, 5 mM HEPES, pH 7.4. Uptake of ³³P_i into oocytes expressing PfPiT was linear with time for at least 60 min (data not shown).

Protein sequence analysis. The phylogenetic tree, hydropathy plot alignment and topographical model were generated as described in Supplementary Methods.

Transfection and immunofluorescence analysis of parasites. Parasites were transfected with the gene encoding a HA-epitope-tagged PfPiT protein (Supplementary Methods). Western blots and immunofluorescence analysis were performed as described previously²².

Statistics. Statistical comparisons were made with either a Student's *t*-test for paired or unpaired samples, or an analysis of variance.

Received 23 June; accepted 8 August 2006.

Published online 27 September 2006.

- Kirk, K. Membrane transport in the malaria-infected erythrocyte. *Physiol. Rev.* **81**, 495–537 (2001).
- Lee, P., Ye, Z., Van Dyke, K. & Kirk, R. G. X-ray microanalysis of *Plasmodium falciparum* and infected red blood cells: effects of qinghaosu and chloroquine on potassium, sodium, and phosphorus composition. *Am. J. Trop. Med. Hyg.* **39**, 157–165 (1988).
- Staines, H. M., Ellory, J. C. & Kirk, K. Perturbation of the pump-leak balance for Na⁺ and K⁺ in malaria-infected erythrocytes. *Am. J. Physiol. Cell Physiol.* **280**, C1576–C1587 (2001).
- Wunsch, S. *et al.* Differential stimulation of the Na⁺/H⁺ exchanger determines chloroquine uptake in *Plasmodium falciparum*. *J. Cell Biol.* **140**, 335–345 (1998).
- Saliba, K. J., Horner, H. A. & Kirk, K. Transport and metabolism of the essential vitamin pantothenic acid in human erythrocytes infected with the malaria parasite *Plasmodium falciparum*. *J. Biol. Chem.* **273**, 10190–10195 (1998).
- Turner, R. J. Stoichiometry of cotransport systems. *Ann. NY Acad. Sci.* **456**, 10–25 (1985).
- Kumler, W. D. & Eiler, J. J. The acid strength of mono and diesters of phosphoric acid. The *n*-alkyl esters from methyl to butyl, the esters of biological importance, and the natural guanidine phosphoric acids. *J. Am. Chem. Soc.* **65**, 2355–2361 (1943).
- Allen, R. J. W. & Kirk, K. The membrane potential of the intraerythrocytic malaria parasite *Plasmodium falciparum*. *J. Biol. Chem.* **279**, 11264–11272 (2004).
- Martin, R. E., Henry, R. I., Abbey, J. L., Clements, J. D. & Kirk, K. The 'permeome' of the malaria parasite: an overview of the membrane transport proteins of *Plasmodium falciparum*. *Genome Biol.* **6**, R26 (2005).
- Gardner, M. J. *et al.* Genome sequence of the human malaria parasite *Plasmodium falciparum*. *Nature* **419**, 498–511 (2002).
- Salaun, C., Rodrigues, P. & Heard, J. M. Transmembrane topology of PiT-2, a phosphate transporter-retrovirus receptor. *J. Virol.* **75**, 5584–5592 (2001).
- Le Roch, K. G. *et al.* Discovery of gene function by expression profiling of the malaria parasite life cycle. *Science* **301**, 1503–1508 (2003).
- Bozdech, Z. *et al.* The transcriptome of the intraerythrocytic developmental cycle of *Plasmodium falciparum*. *PLoS Biol.* **1**, E5 (2003).
- Rager, N., Mamoun, C. B., Carter, N. S., Goldberg, D. E. & Ullman, B. Localization of the *Plasmodium falciparum* PfNT1 nucleoside transporter to the parasite plasma membrane. *J. Biol. Chem.* **276**, 41095–41099 (2001).
- Kavanaugh, M. P. *et al.* Cell-surface receptors for gibbon ape leukemia virus and amphotropic murine retrovirus are inducible sodium-dependent phosphate symporters. *Proc. Natl Acad. Sci. USA* **91**, 7071–7075 (1994).
- Bai, L., Collins, J. F. & Ghishan, F. K. Cloning and characterization of a type III Na-dependent phosphate cotransporter from mouse intestine. *Am. J. Physiol.* **279**, C1135–C1143 (2000).
- Dubyak, G. R. Ion homeostasis, channels, and transporters: an update on cellular mechanisms. *Adv. Physiol. Educ.* **28**, 143–154 (2004).
- Opekarova, M. & Tanner, W. Specific lipid requirements of membrane proteins—a putative bottleneck in heterologous expression. *Biochim. Biophys. Acta* **1610**, 11–22 (2003).
- Bröer, A. *et al.* Molecular cloning of mouse amino acid transport system B⁰, a neutral amino acid transporter related to Hartnup disorder. *J. Biol. Chem.* **279**, 24467–24476 (2004).
- Bröer, S. *et al.* Comparison of lactate transport in astroglial cells and monocarboxylate transporter 1 (MCT 1) expressing *Xenopus laevis* oocytes. Expression of two different monocarboxylate transporters in astroglial cells and neurons. *J. Biol. Chem.* **272**, 30096–30102 (1997).
- Bröer, S. *Xenopus laevis* oocytes. *Methods Mol. Biol.* **227**, 245–258 (2003).
- Mullin, K. A. *et al.* Membrane transporters in the relict plastid of malaria parasites. *Proc. Natl Acad. Sci. USA* **103**, 9572–9577 (2006).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank the Australian Red Cross Blood Service (Canberra and Melbourne) for the provision of blood. This work was supported by grants from the Australian National Health and Medical Research Council (NHMRC), the Australian Research Council (ARC), the Howard Hughes Medical Institute and the ARC/NHMRC Network for Parasitology.

Author Contributions S.B. and K.K. contributed equally to this work and are joint senior authors.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to K.K. (kieran.kirk@anu.edu.au).

LETTERS

Identification of a mammalian mitochondrial porphyrin transporter

Partha C. Krishnamurthy^{1*}, Guoqing Du^{1*}, Yu Fukuda^{1,4}, Daxi Sun¹, Janardhan Sampath¹, Kelly E. Mercer¹, Junfeng Wang², Beatriz Sosa-Pineda², K. Gopal Murti³ & John D. Schuetz¹

The movement of anionic porphyrins (for example, haem) across intracellular membranes is crucial to many biological processes, but their mitochondrial translocation and coordination with haem biosynthesis is not understood. Transport of porphyrins into isolated mitochondria is energy-dependent^{1–3}, as expected for the movement of anions into a negatively charged environment. ATP-binding cassette transporters actively facilitate the transmembrane movement of substances. We found that the mitochondrial ATP-binding cassette transporter ABCB6 is upregulated (messenger RNA and protein in human and mouse cells) by elevation of cellular porphyrins and postulated that ABCB6 has a function in porphyrin transport. We also predicted that ABCB6 is functionally linked to haem biosynthesis, because its mRNA is found in both human bone marrow and CD71⁺ early erythroid cells (by database searching), and because our results show that ABCB6 is highly expressed in human fetal liver, and *Abcb6* in mouse embryonic liver. Here we demonstrate that ABCB6 is uniquely located in the outer mitochondrial membrane and is required for mitochondrial porphyrin uptake. After ABCB6 is upregulated in response to increased intracellular porphyrin, mitochondrial porphyrin uptake activates *de novo* porphyrin biosynthesis. This process is blocked when the *Abcb6* gene is silenced. Our results challenge previous assumptions about the intracellular movement of porphyrins and the factors controlling haem biosynthesis.

The expression of *Abcb6* in mouse and human cell lines, respectively, is not altered (data not shown) by cytoplasmic iron chelation or by ferric ammonium citrate, which is readily taken up by cells. Therefore, *Abcb6* is not directly responsive to cellular iron content despite high expression in human fetal liver and mouse embryonic liver (Fig. 1a, b). However, *Abcb6* mRNA (Supplementary Fig. 1a) and ABCB6 protein (Fig. 1c) are strongly elevated by haemin (ferric haem), which is intracellularly reduced to haem, in murine erythroleukaemia (MEL) cells.

We investigated the relationship of ABCB6 expression to erythroid differentiation in two model systems. We first stimulated erythroid differentiation of MEL cells with dimethyl sulphoxide. As expected, ABCB6 protein increased when haem (used throughout as a generic term denoting no specific valence state) biosynthesis was maximal (not shown). Second, in the proerythroblast cell line G1ER, ABCB6 protein increased coordinately with globin during oestradiol-induced erythroid differentiation (Fig. 1d).

Because erythrocyte maturation and haem biosynthesis are not inextricably linked^{4,5}, we tested whether ABCB6 production is induced by increased intracellular haem. Treatment of MEL cells with amino-levulinic acid (ALA) to induce haem biosynthesis caused a dose- and time-dependent increase in ABCB6 protein and mRNA

(Fig. 2a; Supplementary Fig. 1b, c), and intracellular protoporphyrin IX (PPIX) and ABCB6 concentrations were correlated ($r^2 = 0.9$). Blockade of haem biosynthesis in MEL cells with succinyl acetone⁶ prevented an increase in intracellular PPIX and blocked upregulation of ABCB6 synthesis by ALA (Fig. 2b, c). Therefore, ABCB6 expression levels respond directly to the intracellular PPIX concentration.

In two animal models, we investigated whether *Abcb6* expression responds to altered haem/protoporphyrin levels. In the first model, mice treated with phenylhydrazine to enhance splenic erythropoiesis showed a parallel increase in splenic uroporphyrinogen decarboxylase (a correlate of haem biosynthesis) and *Abcb6* mRNA (Fig. 2d), suggesting that haem biosynthesis induces *Abcb6* *in vivo*. The second model used *Abcg2* (also known as *Bcrp*) knockout mice⁷, which have elevated levels of PPIX^{7,8}. *Abcb6* expression in these mice, as compared with wild-type littermates, was 2-fold greater in primitive bone marrow c-Kit⁺Lin[−] cells (not shown) and 10- to 50-fold greater in heart, intestine and testis (Fig. 2e).

The precise localization of ABCB6 in mitochondria is unknown. We treated K562 cells with haemin and probed the inner-membrane

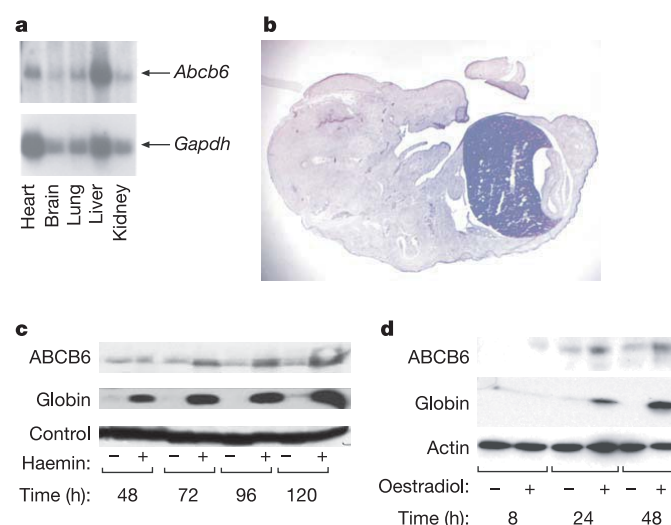


Figure 1 | ABCB6 is highly expressed in human and murine fetal liver, and is upregulated in the presence of haemin and during erythroid differentiation. **a**, Northern blot analysis of ABCB6 in human fetal tissues. **b**, *In situ* hybridization showing abundant hepatic *Abcb6* expression in an mouse embryo at embryonic day E14.5. **c**, Immunoblot analysis of MEL cells with anti-ABCB6 (ref. 13) or anti-globin. 'Control' is a loading control. **d**, Immunoblot analyses in differentiating G1ER cells²⁹ showing expression of ABCB6, and globin protein. Actin is a loading control.

¹Department of Pharmaceutical Sciences and ²Department of Genetics and Tumor Cell Biology, and the ³Scientific Imaging Resource Center, St. Jude Children's Research Hospital, 332 North Lauderdale Street, Memphis, Tennessee 38105-2794, USA. ⁴University of Tennessee Health Sciences Center, Memphis, Tennessee 38103, USA.

*These authors contributed equally to this work.

and outer-membrane mitochondrial fractions for ABCB6, ABCB7, and inner-membrane and outer-membrane markers. ABCB6 was located in the outer membrane and ABCB7 in the inner membrane (Supplementary Fig. 2a). We similarly purified and probed inner membrane and outer membrane from cells ectopically expressing ABCB6–Flag. Like its endogenous counterpart, ABCB6 localized to the outer membrane (Supplementary Fig. 2b). We then used confocal microscopy to assess co-localization of markers in the purified fractions. In the outer membrane, ABCB6–Flag and outer-membrane marker substantially overlapped, and a small amount of inner-membrane marker was observed. The inner-membrane fraction showed no co-localization of ABCB6 and inner-membrane marker but contained a small amount of both ABCB6 and outer-membrane marker (Supplementary Fig. 2c–f).

We next performed a proteolytic clipping assay on mitochondria from cells expressing ABCB6–Flag (Supplementary Fig. 3a). ABCB6 was nearly absent from mitochondria that were isolated after proteinase K digestion, although mitochondrial integrity was preserved (Supplementary Fig. 3b). After removal of the outer mitochondrial membrane by osmotic shock, ABCB6 was not detected in the inner-membrane mitoplasts. These results show for the first time, to our knowledge, that an ATP-binding cassette (ABC) transporter localizes to the outer mitochondrial membrane.

Most eukaryotic ABC transporters are single polypeptides⁹, but a few are homodimeric ‘half-transporters’. Many ABC half-transporters do not form homodimers *in vivo*, and others function only as heterodimers¹⁰. Structural studies of prototypical ABC transporters suggested that ABCB6 forms either a homodimer or a heterodimer. In NIH3T3 cells, we expressed ABCB6 containing either a carboxy-terminal V5 epitope or a Flag tag. Immunoprecipitation with anti-Flag or anti-V5 demonstrated that ABCB6 forms a homodimer (Supplementary Fig. 3c) and is thus unlikely to require a heterodimeric partner.

Localization to the outer mitochondrial membrane would position ABCB6 appropriately for binding and transport of porphyrins. After haemin-agarose affinity chromatography, lysates of cells expressing ABCB6–Flag precipitated as a complex with haemin-agarose⁸. The precipitate contained most of the input ABCB6 (Fig. 3a, top), indicating that ABCB6 is a haem-binding protein. Additional studies showed that membrane-bound ABC transporters do not generally associate with haemin-agarose (not shown). Some cytoplasmic tetrapyrroles, for example, coproporphyrin III (CPIII)^{1–5}, require transport into the mitochondria^{1,4}. Therefore, we investigated which features of pyrroles (including tetrapyrroles such as haem, CPIII, PPIX) are required to displace ABCB6 from haemin-agarose. The monopyrrole porphobilinogen did not displace ABCB6 even at 80 μ M, whereas CPIII, an abundant tetrapyrrole in erythroid cells, readily displaced more than 90% of ABCB6 at 10 μ M (Fig. 3a, bottom). Haemin and

PPIX reduced binding by 50% only at concentrations $>10 \mu$ M. Pheophorbide A, a tetrapyrrole and mitochondrial toxin, was almost as effective as CPIII, whereas cobalamin (vitamin B₁₂), which has a core tetrapyrrole but complex side chains, was ineffective (Fig. 3a, bottom). Therefore, interaction with ABCB6 requires both the anionic carboxylate side chains and the tetrapyrrole moiety.

To test further the specificity of ABCB6 in haem binding, we used a natural ABCB6 splice variant (human P-glycoprotein-related protein, HuPRP) that lacks the transmembrane domain (amino acids 240–289; predicted by <http://SOSUI.proteome.bio.tuat.ac.jp> and the TOPPED program at <http://bioweb.pasteur.fr/seqanal/interfaces/toppred.html>) reported to be the substrate binding site in ABC transporters. HuPRP localized to the outer mitochondrial membrane but did not bind haem (not shown).

ABC transport is fuelled by ATP hydrolysis⁹. We tested the ATP- and time-dependent uptake of ⁵⁵Fe-haemin into mitochondria from cells overexpressing ABCB6. Uptake was time-dependent (Fig. 3b), and initial uptake in the presence of ATP was almost three times greater than that in the presence of AMP (178 versus 66 pmol per min per mg protein; $P < 0.05$, Student's *t*-test). Uptake was negligible at 4 °C, consistent with ATP hydrolysis as the primary energy source. We then expressed a non-functional ABCB6 mutant⁹ and wild-type ABCB6 in K562 cells. Mitochondria from cells expressing wild-type ABCB6, unlike those from cells expressing mutant ABCB6, showed time-dependent uptake of ⁵⁵Fe-haemin (Fig. 3c). Therefore, mitochondrial haem uptake seems to require ATP hydrolysis.

We next investigated whether ABCB6-mediated haemin uptake requires a mono- or tetrapyrrole structure. Haemin uptake by purified mitochondria was not affected by addition of the monopyrrole porphobilinogen, but the tetrapyrrole CPIII reduced uptake in a dose-dependent manner (1 μ M reduced uptake by more than 50%; Fig. 3d). Furthermore, CPIII stimulated vanadate-sensitive ATPase activity only in the mitochondria of cells expressing wild-type ABCB6 (Fig. 3e). The strong competition of CPIII for haemin-agarose binding, potent inhibition of haemin transport, and stimulation of ATPase activity strongly implicate CPIII as an endogenous ABCB6 substrate. Moreover, ABCB6 requires a specific domain for haem binding; the tetrapyrrole ring is essential for substrate binding and transport; and non-functional ABCB6 does not transport haemin or, presumably, other tetrapyrroles.

Because ABCB6 protein increases during erythroid differentiation and haem biosynthesis (Fig. 1d) and is directly upregulated by increased intracellular porphyrin (Fig. 2c), we tested whether ABCB6 expression upregulates porphyrin biosynthesis. Mean PPIX fluorescence values were much greater in ABCB6-overexpressing K562, Saos-2 and MEL cells than in vector-control or ABCB6-mutant cells (Fig. 4a; Supplementary Fig. 4a, b). To determine whether

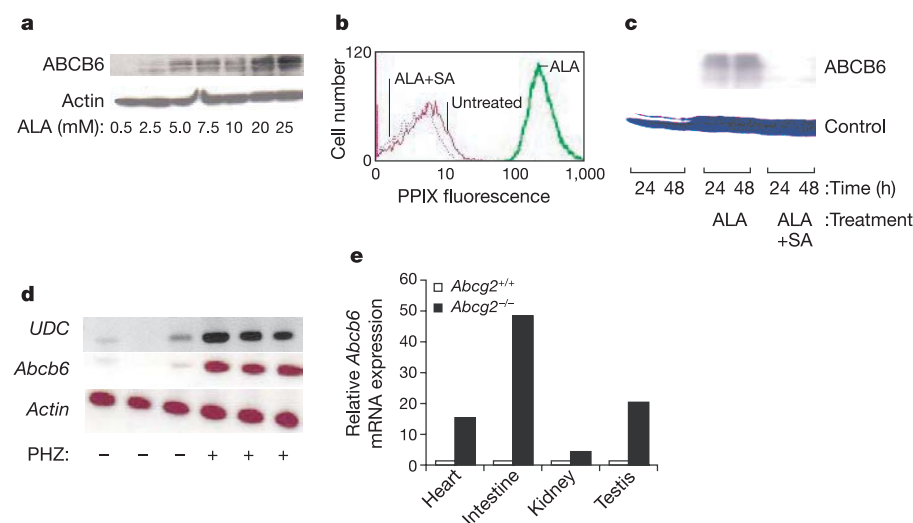


Figure 2 | ABCB6 expression is regulated by intracellular protoporphyrin levels.

a, Immunoblot analysis showing dose-dependent induction of ABCB6 by ALA in MEL cells.

b, Treatment with succinyl acetone (SA) blocks ALA induction of PPIX in MEL cells. **c**, ALA induction of ABCB6 is blocked by SA. **d**, Mice treated with phenylhydrazine (PHZ) show increased splenic expression of *Abcb6* and uroporphyrinogen decarboxylase (UDC; $n = 6$).

e, *Abcb6* mRNA expression in tissues of *Abcg2*^{-/-} mice relative to wild-type mice ($n = 4$ per organ).

increased PPIX reflected accelerated porphyrin biosynthesis, we added radiolabelled glycine (an ALA precursor) to cells and extracted protoporphyrins¹¹. Cells overexpressing ABCB6 had a rate of *de novo* porphyrin biosynthesis 6 to 13 times greater than that of vector-control cells (Fig. 4b, left); this increase was blocked by the ALA-dehydratase inhibitor succinyl acetone⁶ (Fig. 4b, right).

Elevated levels of erythrocyte porphyrin are associated with altered expression of haem biosynthesis enzymes; we assessed the expression in ABCB6-overexpressing cells of gene products that are rate-limiting in erythroid porphyrin biosynthesis. Microarray analyses showed greater expression of transferrin receptor ($\times 3.3$), ALA-dehydratase ($\times 2.0$) and coproporphyrinogen oxidase ($\times 1.5$) RNA in ABCB6-overexpressing MEL cells than in vector-control cells (not shown). In a more detailed real-time polymerase chain reaction analysis

(Supplementary Fig. 5), expression of coproporphyrinogen oxidase ($\times 13$), ALA dehydratase ($\times 7$) and ALA synthase 2 ($\times 7$) was much greater in ABCB6-overexpressing K562 cells than in vector-control cells and proportionately decreased when ABCB6 expression was reduced (not shown). Therefore, ABCB6 overexpression enhances porphyrin biosynthesis by upregulating key genes.

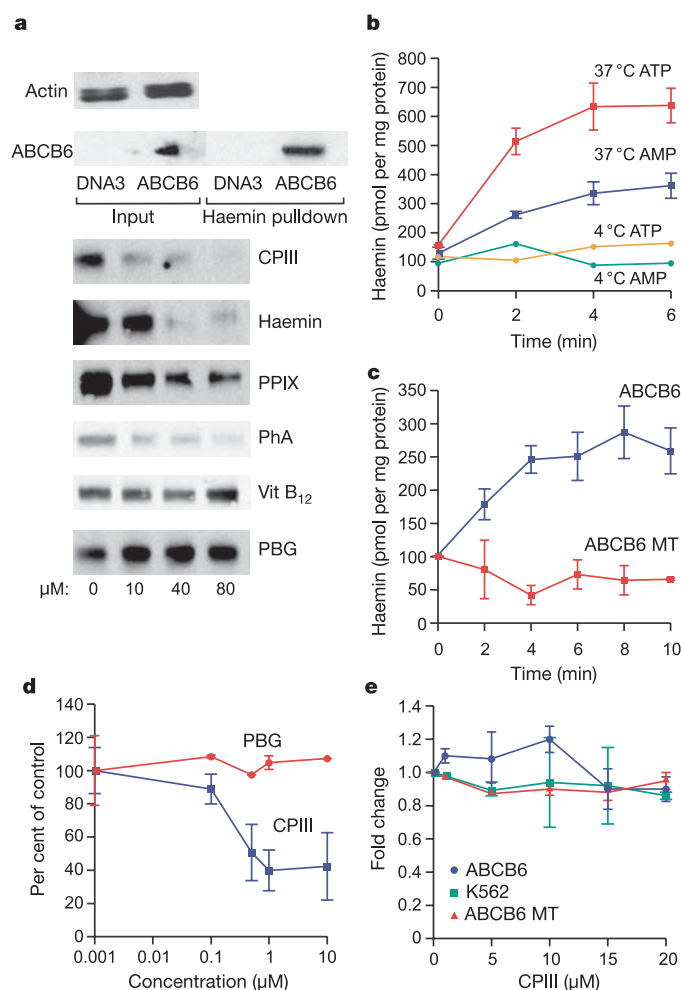


Figure 3 | ABCB6 binds and transports haem. **a**, To test for an interaction between ABCB6 and haem, mitochondria were isolated from Saos-2 cells expressing ABCB6-Flag or empty vector. One portion (input) was analysed by immunoblotting with anti-Flag; another (pull-down) was incubated with haemin-agarose and the resulting complex was also immunoblotted (top two panels). CPIII and haemin potently disrupt the interaction between ABCB6 and haemin-agarose compared with the action of other pyrroles (bottom panels; ABCB6-Flag pull-down probed with anti-Flag). PhA, pheophorbide A. **b**, Uptake of ⁵⁵Fe-labelled haemin by mitochondria from ABCB6-overexpressing cells at 37 °C and 4 °C in the presence of ATP or AMP. **c**, ATP-dependent uptake of ⁵⁵Fe-haemin by mitochondria from K562 cells expressing wild-type or Walker-domain-mutant (MT) ABCB6. **d**, CPIII inhibits mitochondrial haemin uptake in a dose-dependent manner, whereas porphobilinogen (PBG) has no effect. **e**, Vanadate-sensitive ATPase activity (fold change relative to basal activity) was stimulated by CPIII only in mitochondria from cells expressing wild-type ABCB6. Values are means $\pm$ standard error of the mean (s.e.m.) of 2–4 measurements.

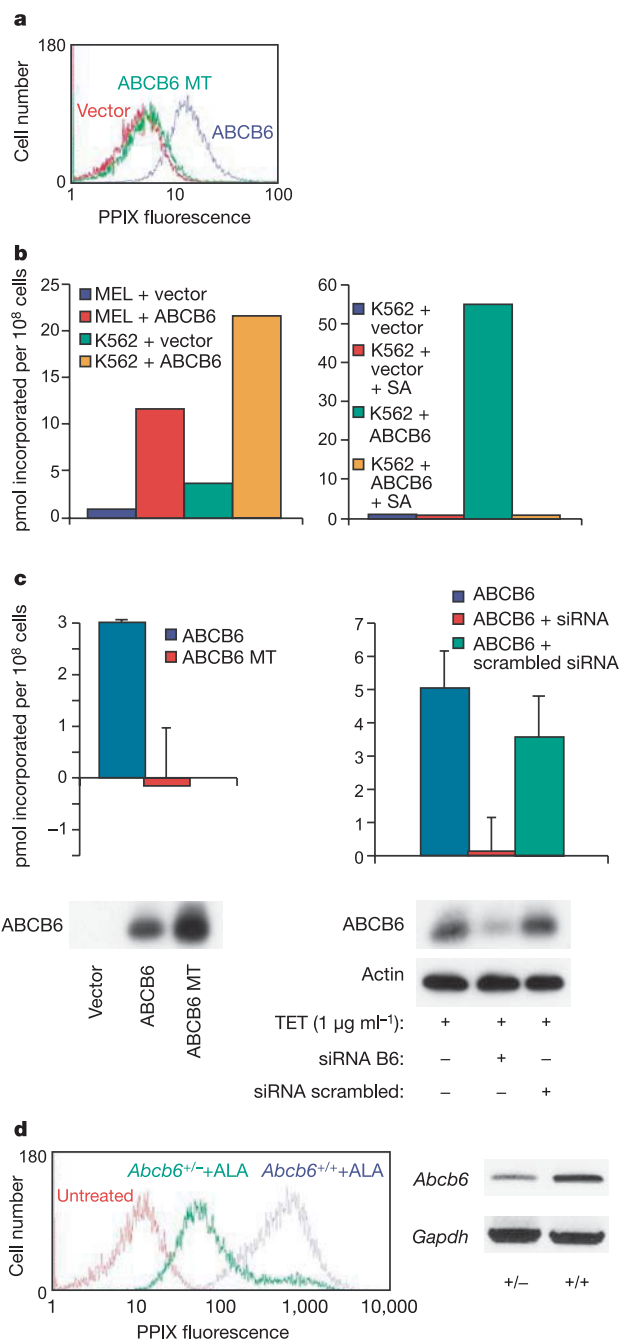


Figure 4 | ABCB6 regulates porphyrin biosynthesis. **a**, Only K562 cells overexpressing wild-type ABCB6 show high PPIX fluorescence signal. **b**, Porphyrin incorporation of ¹⁴C-glycine is increased in ABCB6-overexpressing cells (left), but is blocked by pretreatment with succinyl acetone (SA; right). **c**, Expression of Flag-tagged or Walker-domain-mutant ABCB6 was induced in T-Rex-CHO cells (see Supplementary Methods), and haem biosynthesis was assayed (left; values are the mean $\pm$ s.e.m. of three independent experiments); haem biosynthesis was suppressed by addition of an *Abcb6* siRNA (right; values represent the mean $\pm$ s.e.m. of four independent experiments). TET, tetracycline. **d**, *Abcb6*^{+/-} and *Abcb6*^{+/+} ES cells were treated with ALA, and PPIX was assayed (left); corresponding *Abcb6* mRNA expression of the *Abcb6* genotypes (right).

ABCB6 is expressed in various non-erythroid cells^{12–14} and may have a general role in regulating haem biosynthesis. We transiently expressed ABCB6 or a Walker-A-mutant ABCB6 in Chinese hamster ovary (CHO) cells. Although both proteins were expressed, haem biosynthesis increased only in cells with functional ABCB6 (Fig. 4c, left). Furthermore, addition of short interfering RNA (siRNA) targeting *Abcb6*, but not a scrambled *Abcb6* siRNA, abrogated haem biosynthesis (Fig. 4c, right). To test whether loss of one *Abcb6* allele affects haem biosynthesis, we measured PPIX, haem and *Abcb6* in *Abcb6*^{+/+} and *Abcb6*^{+/-} 129/SVJ-derived embryonic stem (ES) cell lines treated with ALA (see Supplementary Methods). Haem biosynthesis in the *Abcb6*^{+/+} cells was almost twice that in the *Abcb6*^{+/-} cells (6.8 versus 3.8 pmol per 10⁸ cells), as confirmed by PPIX and *Abcb6* expression (Fig. 4d).

We have shown that the mammalian ABC transporter ABCB6 is uniquely associated with the outer mitochondrial membrane and interacts specifically with haem and porphyrins, which it can transport into the mitochondria. CPIII is the most probable endogenous substrate of ABCB6; its transfer to coproporphyrinogen oxidase on the inner mitochondrial membrane¹⁵ would efficiently facilitate haem biosynthesis. Moreover, the coordinated upregulation of ABCB6 and porphyrin biosynthesis would counteract the rate-limiting movement of tetrapyrroles into the mitochondria^{16–21}. We propose that ABCB6 facilitates transmembrane movement of porphyrins into mammalian mitochondria by counteracting kinetic and energetic barriers. This role is consistent with the upregulation of ABCB6 and haem biosynthetic enzymes during differentiation⁴. In non-erythroid cells, ABCB6 porphyrin transport may be involved in the assembly of mitochondrial respiratory-chain enzymes that require haem and in the mitochondrial import of ALA synthase and coproporphyrinogen oxidase, which is blocked by haem²². Finally, mitochondrial compartmentalization of protoporphyrin may protect cells from the degradation of macromolecules (DNA, protein)^{23–26} and the accumulation of plasma membrane protoporphyrin^{25,27,28}.

METHODS

Mitochondrial transport assays. Uptake of Fe⁵⁵-haemin into isolated mitochondria was measured by rapid filtration. Briefly, mitochondrial preparations containing 20 µg of protein were thawed for 30 s at 37 °C in 50 µl TS buffer (50 mM Tris HCl, 250 mM sucrose, pH 7.4). An ATP regenerating mix containing 4 mM ATP or AMP, 10 mM MgCl₂, 10 mM creatine phosphate, 100 µg ml⁻¹ creatine kinase and 23 nM Fe⁵⁵-haemin was then added (final volume, 50 µl) and mitochondria were incubated for the indicated times. The reaction was stopped and immediately filtered through a 0.45-µm cellulose filter. The filter was washed four times with 9 ml of ice-cold TS buffer and analysed by liquid scintillation counting. Fe⁵⁵-haemin uptake was defined as the difference between substrate binding to mitochondria at 4 °C and after incubation at 37 °C. For competition studies, ATP-dependent uptake was compared after a 2-min incubation in the presence or absence of ATP in mitochondria prepared from K562 cells transduced with an ABCB6 expression vector or with an empty vector.

Determination of ABCB6 expression patterns. Database search was performed using *Abcb6* as a keyword (<http://www.genecards.org/index.shtml>, <http://symatlas.gnf.org/SymAtlas/>) to determine the ABCB6 expression patterns in different tissues and cell lineages

Received 25 April; accepted 25 July 2006.

Published online 27 September 2006.

- Rebeiz, N., Arkins, S., Kelley, K. W. & Rebeiz, C. A. Enhancement of coproporphyrinogen III transport into isolated transformed leukocyte mitochondria by ATP. *Arch. Biochem. Biophys.* **333**, 475–481 (1996).
- Koller, M. E. Studies on the uptake of porphyrin by isolated rat liver mitochondria with particular emphasis on the effect of hemin. *FEBS Lett.* **100**, 47–51 (1979).
- Koller, M. E. & Romslo, I. Uptake of protoporphyrin IX by isolated rat liver mitochondria. *Biochem. J.* **188**, 329–335 (1980).
- Ponka, P. Cell biology of heme. *Am. J. Med. Sci.* **318**, 241–256 (1999).
- Tsiftoglou, A. S., Nunez, M. T., Wong, W. & Robinson, S. H. Dissociation of iron transport and heme biosynthesis from commitment to terminal maturation of murine erythroleukemia cells. *Proc. Natl Acad. Sci. USA* **80**, 7528–7532 (1983).

- Sassa, S. & Kappas, A. Succinylacetone inhibits delta-aminolevulinic dehydratase and potentiates the drug and steroid induction of delta-aminolevulinic synthase in liver. *Trans. Assoc. Am. Physicians* **95**, 42–52 (1982).
- Jonker, J. W. *et al.* The breast cancer resistance protein protects against a major chlorophyll-derived dietary phototoxin and protoporphyria. *Proc. Natl Acad. Sci. USA* **99**, 15649–15654 (2002).
- Krishnamurthy, P. *et al.* The stem cell marker Bcrp/ABCG2 enhances hypoxic cell survival through interactions with heme. *J. Biol. Chem.* **279**, 24218–24225 (2004).
- Locher, K. P. Structure and mechanism of ABC transporters. *Curr. Opin. Struct. Biol.* **14**, 426–431 (2004).
- Graf, G. A., Cohen, J. C. & Hobbs, H. H. Missense mutations in ABCG5 and ABCG8 disrupt heterodimerization and trafficking. *J. Biol. Chem.* **279**, 24881–24888 (2004).
- Gardner, L. C. & Cox, T. M. Biosynthesis of heme in immature erythroid cells: the regulatory step for heme formation in the human erythron. *J. Biol. Chem.* **263**, 6676–6682 (1988).
- Mitsuhashi, N. *et al.* MTABC3, a novel mitochondrial ATP-binding cassette protein involved in iron homeostasis. *J. Biol. Chem.* **275**, 17536–17540 (2000).
- Furuya, K. N., Bradley, G., Sun, D., Schuetz, E. G. & Schuetz, J. D. Identification of a new P-glycoprotein-like ATP-binding cassette transporter gene that is overexpressed during hepatocarcinogenesis. *Cancer Res.* **57**, 3708–3716 (1997).
- Hirsch-Ernst, K. I. *et al.* Molecular cDNA cloning and tissue distribution of mRNA encoding a novel ATP-binding cassette (ABC) half-transporter. *Biochem. Biophys. Res. Commun.* **249**, 151–155 (1998).
- Ferreira, G. C., Andrew, T. L., Karr, S. W. & Dailey, H. A. Organization of the terminal two enzymes of the heme biosynthetic pathway. Orientation of protoporphyrinogen oxidase and evidence for a membrane complex. *J. Biol. Chem.* **263**, 3835–3839 (1988).
- Light, W. R. III & Olson, J. S. The effects of lipid composition on the rate and extent of heme binding to membranes. *J. Biol. Chem.* **265**, 15632–15637 (1990).
- Light, W. R. III & Olson, J. S. Transmembrane movement of heme. *J. Biol. Chem.* **265**, 15623–15631 (1990).
- Ricchelli, F., Gobbo, S., Jori, G., Salet, C. & Moreno, G. Temperature-induced changes in fluorescence properties as a probe of porphyrin microenvironment in lipid membranes. 2. The partition of hematoporphyrin and protoporphyrin in mitochondria. *Eur. J. Biochem.* **233**, 165–170 (1995).
- Lavi, A., Weiman, H., Holmes, R. T., Smith, K. M. & Ehrenberg, B. The depth of porphyrin in a membrane and the membrane's physical properties affect the photosensitizing efficiency. *Biophys. J.* **82**, 2101–2110 (2002).
- Kepczynski, M., Pandian, R. P., Smith, K. M. & Ehrenberg, B. Do liposome-binding constants of porphyrins correlate with their measured and predicted partitioning between octanol and water? *Photochem. Photobiol.* **76**, 127–134 (2002).
- Kuzelova, K. & Brault, D. Kinetic and equilibrium studies of porphyrin interactions with unilamellar lipidic vesicles. *Biochemistry* **33**, 9447–9459 (1994).
- Susa, S. *et al.* Heme inhibits the mitochondrial import of coproporphyrinogen oxidase. *Blood* **100**, 4678–4679 (2002).
- Aft, R. L. & Mueller, G. C. Hemin-mediated DNA strand scission. *J. Biol. Chem.* **258**, 12069–12072 (1983).
- Aft, R. L. & Mueller, G. C. Hemin-mediated oxidative degradation of proteins. *J. Biol. Chem.* **259**, 301–305 (1984).
- Kuross, S. A., Rank, B. H. & Hebbel, R. P. Excess heme in sickle erythrocyte inside-out membranes: possible role in thiol oxidation. *Blood* **71**, 876–882 (1988).
- Krieg, R. C. *et al.* Intracellular localization is a cofactor for the phototoxicity of protoporphyrin IX in the gastrointestinal tract: *in vitro* study. *Photochem. Photobiol.* **78**, 393–399 (2003).
- Piomelli, S., Brickman, A. & Carlos, E. Rapid diagnosis of iron deficiency by measurement of free erythrocyte porphyrins and hemoglobin: the FEP/hemoglobin ratio. *Pediatrics* **57**, 136–141 (1976).
- Sassa, S. & Kappas, A. Molecular aspects of the inherited porphyrias. *J. Intern. Med.* **247**, 169–178 (2000).
- Visvader, J. E., Mao, X., Fujiwara, Y., Hahm, K. & Orkin, S. H. The LIM-domain binding protein Ldb1 and its partner LMO2 act as negative regulators of erythroid differentiation. *Proc. Natl Acad. Sci. USA* **94**, 13707–13712 (1997).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank S. Naron for editorial assistance; J. Groff for preparation of the illustrations; A. Smith for comments and suggestions; S. Orkin for providing the G1ER cells; B. Sarkadi and C. Ozvegy-Laczka for assistance with the ATPase assays; and Y. Fan for bioinformatics analysis of *Abcb6*. This work was supported by NIH grants and by the American Lebanese Syrian Associated Charities (ALSAC).

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to J.D.S. (john.schuetz@stjude.org).

LETTERS

Molecular architecture and assembly of the DDB1-CUL4A ubiquitin ligase machinery

Stephane Angers^{1,2,*†}, Ti Li^{2,*}, Xianhua Yi², Michael J. MacCoss³, Randall T. Moon^{1,2,4} & Ning Zheng²

Protein ubiquitination is a common form of post-translational modification that regulates a broad spectrum of protein substrates in diverse cellular pathways¹. Through a three-enzyme (E1–E2–E3) cascade, the attachment of ubiquitin to proteins is catalysed by the E3 ubiquitin ligase, which is best represented by the superfamily of the cullin-RING complexes^{2,3}. Conserved from yeast to human, the DDB1-CUL4A-ROC1 complex is a recently identified cullin-RING ubiquitin ligase, which regulates DNA repair^{4–10}, DNA replication^{11–14} and transcription¹⁵, and can also be subverted by pathogenic viruses to benefit viral infection¹⁶. Lacking a canonical SKP1-like cullin adaptor and a defined substrate recruitment module, how the DDB1-CUL4A-ROC1 E3 apparatus is assembled for ubiquitinating various substrates remains unclear. Here we present crystallographic analyses of the virally hijacked form of the human DDB1-CUL4A-ROC1 machinery, which show that DDB1 uses one β -propeller domain for cullin scaffold binding and a variably attached separate double- β -propeller fold for substrate presentation. Through tandem-affinity purification of human DDB1 and CUL4A complexes followed by mass spectrometry analysis, we then identify a novel family of WD40-repeat proteins, which directly bind to the double-propeller fold of DDB1 and serve as the substrate-recruiting module of the E3. Together, our structural and proteomic results reveal the structural mechanisms and molecular logic underlying the assembly and versatility of a new family of cullin-RING E3 complexes.

Distinct from all known cullin adaptors, DDB1 is a large multi-domain protein featuring an independent β -propeller (BP) domain (BPB) flexibly connected to a clam-shaped double-propeller fold (BPA–BPC)¹⁷. To reveal how DDB1 incorporates into the CUL4A–ROC1 complex and mediates substrate recruitment, we determined the 3.1 Å crystal structure of a DDB1-CUL4A-ROC1 complex bound to the V protein of simian virus 5 (SV5-V), previously known as the VDC complex (Table 1). By docking the STAT proteins to DDB1, the V protein of the paramyxovirus is able to accelerate the ubiquitination

and degradation of STAT1, thereby blocking the antiviral response of the host¹⁶. With all the components in their full-length forms, the 240 kDa quaternary protein complex represents a functionally complete form of the E3 apparatus.

The DDB1-CUL4A-ROC1-SV5-V complex has an overall elongated shape that is about 170 Å long and 100 Å wide, resembling SCF (SKP1-cullin-F-box-protein)¹⁸ (Fig. 1). As a close paralogue of CUL1, CUL4A adopts a CUL1-like structure with an arc-shaped

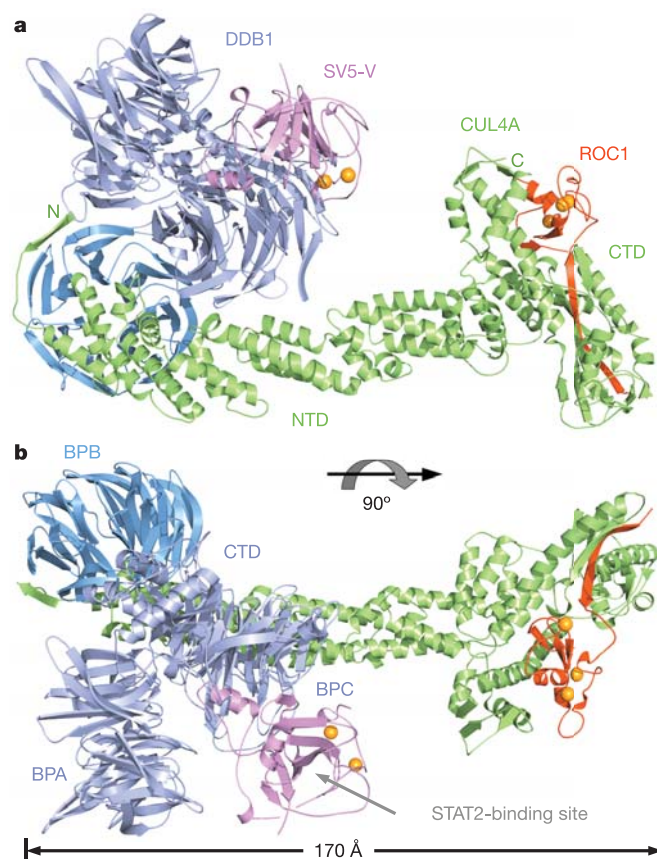


Figure 1 | Crystal structure of the DDB1-CUL4A-ROC1 ubiquitin ligase complex hijacked by the V protein of simian virus 5. **a, b**, Two orthogonal views of the complex structure are shown in ribbon diagram. DDB1, CUL4A, ROC1 and SV5-V are coloured in blue, green, red and magenta, respectively. The BPB domain of DDB1 is coloured in cyan. The zinc atoms in ROC1 and SV5-V are represented as orange spheres. A previously identified STAT2-binding site¹⁷ of the viral protein, as well as the four domains of DDB1, are indicated.

Table 1 | X-ray refinement statistics

Resolution (Å)	3.1
R_{work} ; R_{free} (%)	25.1; 31.6
Number of atoms	
Protein	16,937
Ligand/ion	5
Water	0
B-factors	
Protein	75.9
Ligand/ion	93.5
Water	NA
Root-mean-squared deviations	
Bond lengths (Å)	0.008
Bond angles (°)	1.57

¹Howard Hughes Medical Institute, ²Department of Pharmacology, ³Department of Genome Sciences, ⁴Institute for Stem Cell and Regenerative Medicine, University of Washington, School of Medicine, Box 357280, Seattle, Washington 98195, USA. [†]Present address: Leslie Dan Faculty of Pharmacy, University of Toronto, Toronto, Ontario M5S 3M2, Canada.

*These authors contributed equally to this work.

helical amino-terminal domain (NTD) and a globular carboxy-terminal domain (CTD). Lacking the SKP1/BTB fold, DDB1 devotes its entire BPB domain for CUL4A binding, anchoring at the tip of the NTD of cullin. The BPA–BPC double-propeller fold of DDB1, however, is positioned in the complex facing towards the E2-binding subunit of the E3 ligase, ROC1. As a result, the viral protein—which sits at the opening of the large pocket of the DDB1 double-propeller fold—is about 45 Å away from the RING finger subunit. By comparing this overall architecture of the virally hijacked DDB1–CUL4A–ROC1 complex with SCF¹⁸, we conclude that the cullin-docking function and the substrate-presenting function of DDB1 are separately performed by the BPB domain and the BPA–BPC double-propeller fold, respectively.

The complex formation between DDB1 and CUL4A is predominantly mediated through two separate interfaces, both involving the BPB domain of DDB1 (Fig. 2a). At one interface, the top surface of the propeller fold of DDB1–BPB is recognized by the H2 and H5 helices of CUL4A, which are unexpectedly the same ones used by CUL1 to bind SKP1¹⁸ (compare Fig. 2b, c). At the other interface, CUL4A further embraces the BPB domain of DDB1 using an N-terminal extension sequence, which is unique among all cullin family members¹⁷. Running across two blades of DDB1–BPB, this CUL4A N-terminal sequence wraps around the propeller of DDB1 at the peripheral side via a hydrogen-bond network and hydrophobic packing (Fig. 2a, d). The equal importance of these two separate interfaces in DDB1–CUL4A binding is reflected in the strict conservation of most of the interacting residues among CUL4 and DDB1 orthologues (Supplementary Fig. 1). Furthermore, mutations of the DDB1 or CUL4A residues involved in either interface can significantly weaken the complex formation^{13,17}. Overall, the bipartite binding mode between DDB1 and CUL4A not only strongly supports a specific role of DDB1 in the CUL4A-based E3 machinery, but also demonstrates a surprising variation in cullin–adaptor interactions.

The crystal structures of DDB1 alone and in complex with SV5-V have previously revealed two distinct structural states of DDB1¹⁷, which show dramatic differences in the relative orientation between the substrate-presenting BPA–BPC double-propeller fold and the cullin-binding BPB domain. In the quaternary-complex crystal, DDB1 adopts yet a new domain arrangement. Intriguingly, super-

position analysis of the three structures indicates that all three domain configurations of DDB1 are in fact geometrically allowed in the context of the full DDB1–CUL4A–ROC1 complex (Fig. 2e). The BPA–BPC double-propeller folds of DDB1 in the resulting three models all similarly point towards the E2-binding RING subunit, even though the orientations of the fold itself are correlated by large degrees of rotation (compare Figs 1b and 2e). Because the BPA–BPC double-propeller fold of DDB1 has a very limited number of contacts with the cullin scaffold (Supplementary Fig. 2), we speculate that on binding to different cellular factors, DDB1 might be able to undergo conformational changes, which can modulate the ubiquitin ligase activity of the E3 machinery.

Although a multidomain structure might allow DDB1 to directly recruit substrates to the E3 platform, ubiquitination of several known DDB1–CUL4A substrates has been shown to require additional cellular factors^{7,19}. Mimicked by the viral V protein, endogenous substrate-specific receptors might be used by DDB1 to expand its substrate repertoire. To further clarify the substrate targeting mechanisms of the DDB1–CUL4A E3 machinery, we used a proteomic approach optimized for isolating CUL4A and DDB1 protein complexes from mammalian cells by tandem-affinity purification²⁰. In these analyses, more than 30 proteins were consistently identified by mass spectrometry, giving rise to a protein–interaction network shown in Fig. 3a.

Notably, besides several known cullin regulators and DDB1-binding proteins, numerous novel protein factors were identified in both the CUL4A and DDB1 complexes. In particular, a group of 16 proteins, including DDB2 and CSA, clearly stands out by sharing a predicted WD40 domain. The consistent presence and abundance of these WD40 proteins in the complexes strongly suggest that they represent bona fide members of the human DDB1–CUL4A E3 machinery (Supplementary Tables 2 and 3). It is also of note that reciprocal tandem-affinity pull-down analyses with DDB2 and two randomly selected WD40 proteins (H326 and WDTC1) confirmed that they are individually associated with the DDB1–CUL4A complex (Fig. 3a). Because many of these WD40 proteins do not have any known function other than DDB1–CUL4A binding, we propose to name them the DCAF (DDB1–CUL4A-associated factor) proteins (Supplementary Table 2).

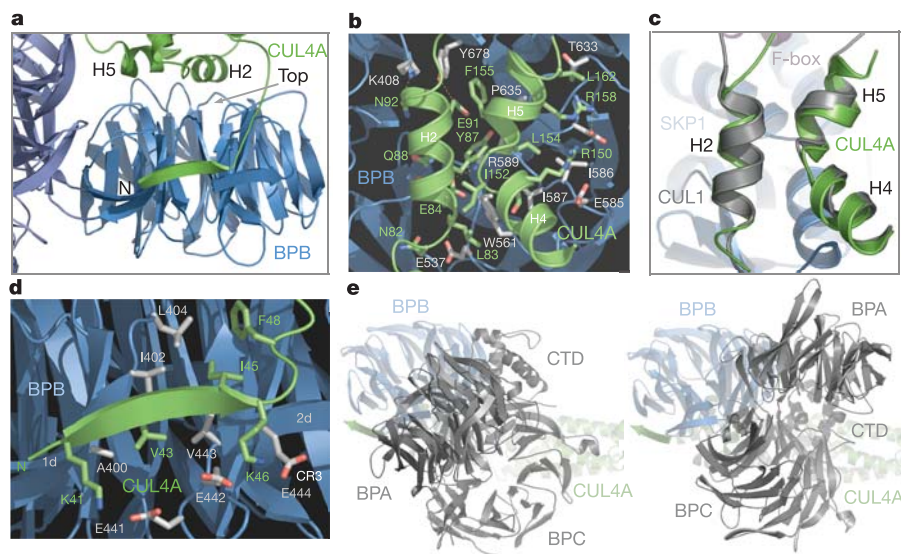


Figure 2 | Binding interfaces and structural flexibility of DDB1 on CUL4A. **a**, An overall view of the structural elements mediating the binding of the DDB1 BPB domain (cyan) with the CUL4A NTD (green). **b**, The interface between the top surface of DDB1–BPB (cyan) and the H2–H5 helices of CUL4A (green) viewed from above with participating side chains shown as sticks (grey in DDB1 and green in CUL4A). **c**, Superposition of the H2–H5 helices of CUL4A (green) and CUL1 (grey) with CUL1-bound SKP1 (blue) shown in the background. The view is the same as that shown in panel **b**. CUL4A-bound DDB1 is not shown for clarity. **d**, The interface between the CUL4A N-terminal extension sequence (green) and the peripheral side of the DDB1 BPB domain (blue). **e**, Superposition analyses of DDB1 alone (right) and DDB1–SV5-V complex (left) structures in the context of the full DDB1–CUL4A–ROC1 complex. The models are docked through the BPB domain and viewed from the same angles as shown in Fig. 1b.

As DDB2 and CSA have already been demonstrated necessary for DDB1–CUL4A-mediated ubiquitination of certain substrates, we speculated that the DCAF proteins we have identified might represent a novel family of proteins serving as the substrate-recruiting module for the E3 machinery. Using co-affinity purification experiments, we first confirmed our mass spectrometry results by showing the *in vivo* interactions of DDB1 with a panel of randomly selected DCAF proteins (Supplementary Fig. 3). Given that both DDB2 and CSA have been reported to directly bind DDB1, we next examined the direct interactions between DDB1 and two DCAF proteins. As shown in Fig. 3b, bacterially produced and purified glutathione S-transferase (GST)-tagged WDR23 (DCAF11) and H326 (DCAF8) were indeed able to interact with the purified recombinant DDB1 protein in an *in vitro* GST pull-down experiment.

If the DCAF proteins represent a family of substrate receptors for DDB1–CUL4A, then they might share a common DDB1-docking motif, analogous to the substrate receptors of other cullin E3 complexes³. As no recognizable common sequence other than the WD40 repeats was found among the DCAF proteins, we postulated that this motif might consist of amino acids exposed on the surface of their WD40 domains. Structure-based sequence analysis of the DCAF proteins revealed the presence of two conserved DxR motifs within the WD40 domain of almost all members (Fig. 3c). These two motifs are separately located at the end of strand c of two consecutive blades of the WD40 fold, and we propose to name these two motifs the 'double DxR box'. On the basis of several known WD40 domain structures^{21,22}, the conserved Asp and Arg residues in the double DxR

box are predicted to be solvent-exposed, lining at the bottom surface of the WD40 propeller fold (Supplementary Fig. 4).

Although other unique and hidden features of the DCAF proteins might contribute to their specific interactions with DDB1, several lines of evidence suggest that the double DxR box is one of the main determinants for their binding to the CUL4A adaptor. First, a naturally occurring mutation (R273H) found in a subset of xeroderma pigmentosum group E (XPE) patients has been mapped to an arginine within the double DxR box motif present in DDB2. As previously reported²³ and confirmed by us (Fig. 3d, lane 2), this mutation abolishes the binding of DDB2 to DDB1. Second, our mutagenesis analysis of two representative DCAF proteins, WDR23 (DCAF11) and H326 (DCAF8), shows that both arginine residues in the double DxR box contribute to DDB1 binding. As shown in Fig. 3d, although single mutation of either of the two arginine residues in the double DxR box of the two DCAF proteins markedly weakens DDB1 binding (lanes 4, 5 and 8, 9), the double mutations completely ablate complex formation (lanes 6 and 10). Third, the double DxR box in two consecutive WD40 blades are also found in two fission yeast WD40 proteins, Cdt2 and Dos1 (also known as Ctr8) (Fig. 3c), which are known to function with yeast DDB1–Cul4, and its close relative Rik1–Cul4, E3 complexes^{19,24,25}. The double DxR motif, therefore, probably underlies an evolutionarily conserved mechanism for docking the substrate-recruiting module to the DDB1–CUL4A ubiquitin ligase machinery.

The presence of a conserved motif shared among the DCAF proteins predicts a common docking site on DDB1. The overall

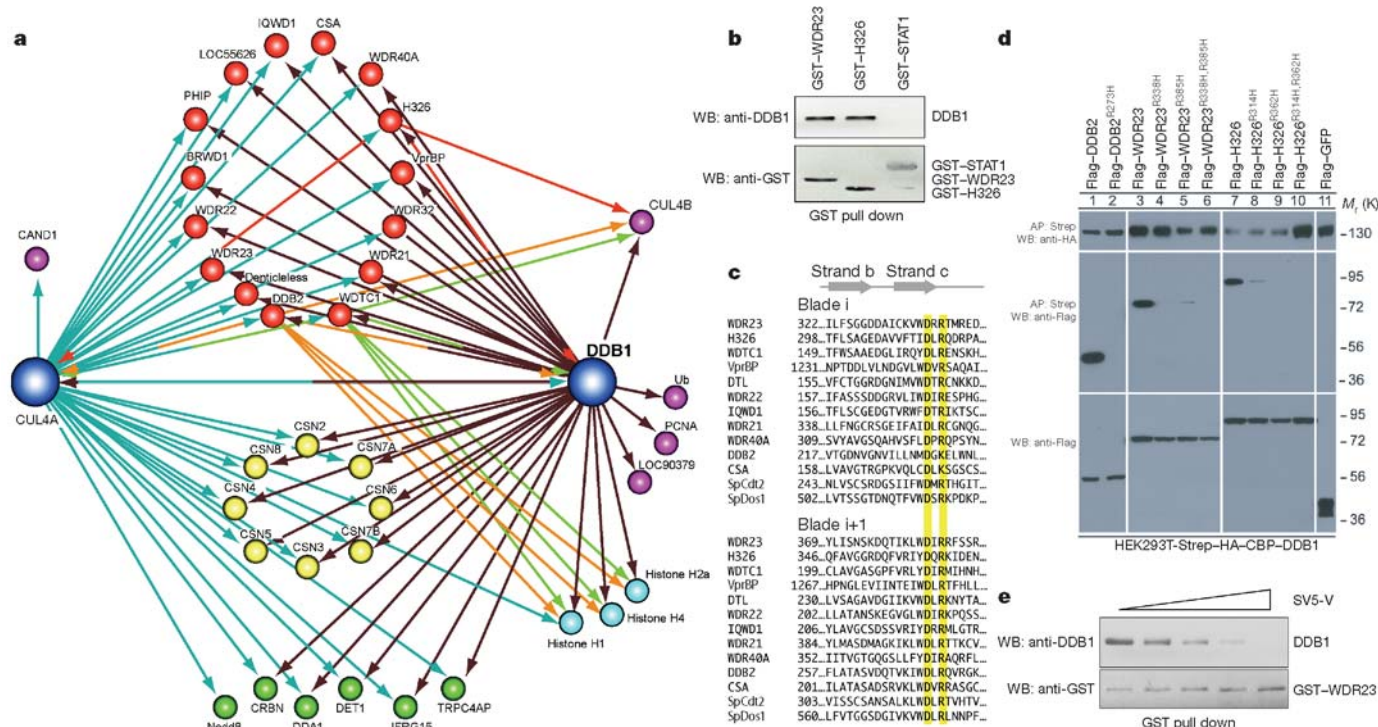


Figure 3 | Identification of a novel family of WD40 proteins as the potential substrate receptors for the DDB1–CUL4A E3 machinery. **a**, Mammalian protein interaction network of the DDB1–CUL4A ubiquitin ligases established by tandem-affinity analyses of CUL4A, DDB1, DDB2, WDTC1 and H326. The novel family of DDB1–CUL4A-associated WD40 domain proteins (DCAF proteins) is shown in red. Arrows represent unidirectional (single arrowhead) and bidirectional (double arrowhead) interactions. **b**, *In vitro* GST pull-down assay followed by western blot (WB) analysis showing the direct interactions between DDB1 and two DCAF proteins, WDR23 (DCAF11) and H326 (DCAF8), with STAT1 as a negative control. **c**, Sequence alignment showing the DxR motifs (highlighted in yellow) conserved in two consecutive β -propeller blades of the DCAF proteins identified in our studies and two yeast WD40 proteins known to bind the

DDB1 (Rik1)–Cul4A complex. The two β -strands preceding the DxR motifs are indicated. **d**, Mutagenesis of the arginine residues in the double DxR box of two representative DCAF proteins impairs DDB1 binding. Interactions between DDB1 and the wild-type and mutant DCAF proteins were assessed by co-affinity purification (AP) using streptavidin (DDB1 pull down) and western blot (WB) analysis using anti-Flag antibodies (DCAF proteins) (middle panel, lanes 1–10). DDB1 purification efficiency and the expression of the DCAF proteins were monitored as shown in the top and bottom panels, respectively. Flag–GFP was used as a negative control. **e**, Competition between SV5-V and WDR23 (DCAF11) for DDB1 binding. The experiments shown in panel **b** were repeated with increasing amounts of purified SV5-V protein. A molar ratio of 1:2 (DDB1:SV5-V) was reached in the final lane.

architecture of the virally hijacked DDB1–CUL4A–ROC1 complex strongly suggests that such a site might be located on the BPA–BPC double-propeller fold of DDB1, probably occupied by SV5-V, even though the viral protein does not have a WD40 domain. In fact, in an *in vitro* binding assay, association of DDB1 and two representative DCAF proteins, WDR23 (DCAF11) and WDTC1 (DCAF9) (data not shown), can be stoichiometrically competed off by increasing amounts of purified SV5-V protein (Fig. 3e). This *in vitro* result is consistent with a previous report showing that the same viral protein can compete with DDB2 for binding with DDB1²⁶. It is interesting to note that, in order to anchor itself to DDB1, the viral protein not only binds to the surface of the BPA–BPC double-propeller fold of DDB1, but also inserts an entire helix into the large pocket of the DDB1 double-propeller fold. Although the size and secondary structure elements of a regular WD40 domain do not allow it to do so, it remains possible that some DCAF proteins might use sequence regions outside their WD40 domains to bind to the pocket of the DDB1 double-propeller fold, thereby further securing their associations with the CUL4A adaptor. Such an assembly mode could be used by the few DCAF proteins that lack a perfect double DxR box motif. Exactly how the DCAF proteins are docked to DDB1 awaits future structural studies.

As gene expression varies in different cell types and under different cellular states, the DCAF proteins identified in our studies are not expected to represent the complete family of substrate receptors of the DDB1–CUL4A E3 in human cells. Having established the structural and molecular basis of the assembly of the E3 machinery, our studies pave the road for a future understanding of the functions of the DDB1–CUL4A complex in eukaryotic biology and in human diseases.

METHODS

Protein preparation. Human CUL4A and ROC1 were co-expressed and purified from *Escherichia coli* and mixed with previously purified DDB1 and SV5-V¹⁷. WDR23 (DCAF11) (amino-acid residues 39–546) and H326 (DCAF8) (amino-acid residues 177–555) were overexpressed and purified from *E. coli* as GST-fusion proteins.

Crystallography. The VDC-complex crystals were grown at 4 °C by the hanging-drop diffusion method. The crystals grew from 100 mM NaHEPES, 7–9% (w/v) PEG 4000, 10% (v/v) isopropanol, 5 mM dithiothreitol, pH 8.0. A native data set was collected at the 8.2.2 beamline at the Advanced Light Source in Berkeley, California, using a crystal flash-frozen in the crystallization buffer supplemented with 20% (v/v) 2-methyl-2,4-pentanediol (MPD) at –170 °C. The crystal forms in space group C222₁, with *a* = 83.46 Å, *b* = 203.19 Å, *c* = 424.87 Å, and one complex in the asymmetric unit. For further details, see Supplementary Methods and Supplementary Fig. 5.

Tandem-affinity purification. Tandem-affinity purification (TAP) and mass spectrometry analyses of the DDB1 and CUL4A protein complexes were performed using HEK293T cells stably expressing the TAP-tagged constructs. The detailed protocol has been described previously²⁰.

Co-affinity purification and GST pull-down assay. Co-affinity purifications were performed using HEK293T cells stably expressing pGLUE-DDB1. Cells were transfected with different Flag-tagged DCAF proteins using CaPO₄. Forty-eight hours post-transfection, cells were lysed using stringent radio-immunoprecipitation buffer (RIPA; 25 mM Tris-HCl, pH 8.0, 250 mM NaCl, 10% (v/v) glycerol, 1% (v/v) Igepal CA-630, 0.25% (w/v) deoxycholic acid, 2 mM EDTA, 1 mM NaF, 50 mM β-glycerophosphate). Lysates were cleared by centrifugation and affinity purification was performed using sepharose–streptavidin resin (Amersham). Anti-Flag M2 antibody (F1804; Sigma) and anti-haemagglutinin (HA) antibody (3F10; Roche) were used to detect the Flag-tagged DCAF proteins and DDB1, respectively. Previously reported procedures¹⁸ were followed for the *in vitro* GST pull-down experiments except that the final samples were analysed by western blot analysis using an anti-DDB1 antibody (ab13562; abcam).

Received 15 June; accepted 17 August 2006.

Published online 10 September 2006.

- Hershko, A. & Ciechanover, A. The ubiquitin system. *Annu. Rev. Biochem.* **67**, 425–479 (1998).
- Pickart, C. M. Mechanisms underlying ubiquitination. *Annu. Rev. Biochem.* **70**, 503–533 (2001).
- Petroski, M. D. & Deshaies, R. J. Function and regulation of cullin-RING ubiquitin ligases. *Nature Rev. Mol. Cell Biol.* **6**, 9–20 (2005).

- Nag, A., Bondar, T., Shiv, S. & Raychaudhuri, P. The xeroderma pigmentosum group E gene product DDB2 is a specific target of cullin 4A in mammalian cells. *Mol. Cell. Biol.* **21**, 6738–6747 (2001).
- Chen, X., Zhang, Y., Douglas, L. & Zhou, P. UV-damaged DNA-binding proteins are targets of CUL-4A-mediated ubiquitination and degradation. *J. Biol. Chem.* **276**, 48175–48182 (2001).
- Sugasawa, K. *et al.* UV-induced ubiquitylation of XPC protein mediated by UV-DDB-ubiquitin ligase complex. *Cell* **121**, 387–400 (2005).
- Groisman, R. *et al.* CSA-dependent degradation of CSB by the ubiquitin–proteasome pathway establishes a link between complementation factors of the Cockayne syndrome. *Genes Dev.* **20**, 1429–1434 (2006).
- Groisman, R. *et al.* The ubiquitin ligase activity in the DDB2 and CSA complexes is differentially regulated by the COP9 signalosome in response to DNA damage. *Cell* **113**, 357–367 (2003).
- Kapetanaki, M. G. *et al.* The DDB1–CUL4A^{DDB2} ubiquitin ligase is deficient in xeroderma pigmentosum group E and targets histone H2A at UV-damaged DNA sites. *Proc. Natl Acad. Sci. USA* **103**, 2588–2593 (2006).
- Wang, H. *et al.* Histone H3 and H4 ubiquitylation by the CUL4–DDB–ROC1 ubiquitin ligase facilitates cellular response to DNA damage. *Mol. Cell* **22**, 383–394 (2006).
- Liu, C. *et al.* Cop9/signalosome subunits and Pcu4 regulate ribonucleotide reductase by both checkpoint-dependent and -independent mechanisms. *Genes Dev.* **17**, 1130–1140 (2003).
- Higa, L. A., Mihaylov, I. S., Banks, D. P., Zheng, J. & Zhang, H. Radiation-mediated proteolysis of CDT1 by CUL4–ROC1 and CSN complexes constitutes a new checkpoint. *Nature Cell Biol.* **5**, 1008–1015 (2003).
- Hu, J., McCall, C. M., Ohta, T. & Xiong, Y. Targeted ubiquitination of CDT1 by the DDB1–CUL4A–ROC1 ligase in response to DNA damage. *Nature Cell Biol.* **6**, 1003–1009 (2004).
- Bondar, T., Ponomarev, A. & Raychaudhuri, P. Ddb1 is required for the proteolysis of the *Schizosaccharomyces pombe* replication inhibitor Spd1 during S phase and after DNA damage. *J. Biol. Chem.* **279**, 9937–9943 (2004).
- Wertz, I. E. *et al.* Human De-etiolated-1 regulates c-Jun by assembling a CUL4A ubiquitin ligase. *Science* **303**, 1371–1374 (2004).
- Horvath, C. M. Weapons of STAT destruction. Interferon evasion by paramyxovirus V protein. *Eur. J. Biochem.* **271**, 4621–4628 (2004).
- Li, T., Chen, X., Garbutt, K. C., Zhou, P. & Zheng, N. Structure of DDB1 in complex with a paramyxovirus V protein: viral hijack of a propeller cluster in ubiquitin ligase. *Cell* **124**, 105–117 (2006).
- Zheng, N. *et al.* Structure of the Cul1–Rbx1–Sklp1–F box^{Skp2} SCF ubiquitin ligase complex. *Nature* **416**, 703–709 (2002).
- Liu, C. *et al.* Transactivation of *Schizosaccharomyces pombe* *cdt2*⁺ stimulates a Pcu4–Ddb1–CSN ubiquitin ligase. *EMBO J.* **24**, 3940–3951 (2005).
- Angers, S. *et al.* The KLHL12–Cullin-3 ubiquitin ligase negatively regulates the Wnt–β-catenin pathway by targeting Dishevelled for degradation. *Nature Cell Biol.* **8**, 348–357 (2006).
- Lambright, D. G. *et al.* The 2.0 Å crystal structure of a heterotrimeric G protein. *Nature* **379**, 311–319 (1996).
- Wall, M. A. *et al.* The structure of the G protein heterotrimer G_{12i}βγ₂. *Cell* **83**, 1047–1058 (1995).
- Shiyanov, P. *et al.* The naturally occurring mutants of DDB are impaired in stimulating nuclear import of the p125 subunit and E2F1-activated transcription. *Mol. Cell. Biol.* **19**, 4935–4943 (1999).
- Li, F. *et al.* Two novel proteins, dos1 and dos2, interact with rik1 to regulate heterochromatic RNA interference and histone modification. *Curr. Biol.* **15**, 1448–1457 (2005).
- Thon, G. *et al.* The Clr7 and Clr8 directionality factors and the Pcu4 cullin mediate heterochromatin formation in the fission yeast *Schizosaccharomyces pombe*. *Genetics* **171**, 1583–1595 (2005).
- Leupin, O., Bontron, S. & Strubin, M. Hepatitis B virus X protein and simian virus 5 V protein exhibit similar UV-DDB1 binding properties to mediate distinct activities. *J. Virol.* **77**, 6274–6283 (2003).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank the beamline staff of the Advanced Light Source at Berkeley for help with data collection, P. Zhou, W. Xu, J. Beavo, Z. Gao and members of the Zheng and Moon laboratories for discussions and help. We also thank J. W. Harper and Y. Xiong for communicating unpublished results. This work is supported by grants from the National Institutes of Health (to N.Z. and M.J.M.), the Pew Scholar Program (N.Z.), and the Howard Hughes Medical Institute (S.A. and R.T.M.).

Author Contributions S.A. and T.L. contributed equally to this work.

Author Information Structure coordinates for the DDB1–CUL4A–ROC1–SV5-V complex have been deposited in the Protein Data Bank under accession number 2HYE. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to N.Z. (nzheng@u.washington.edu) or R.T.M. (rtmoon@u.washington.edu).

LETTERS

Co-evolution of transcriptional and post-translational cell-cycle regulation

Lars Juhl Jensen^{1*}, Thomas Skøt Jensen^{2*}, Ulrik de Lichtenberg^{2*}, Søren Brunak² & Peer Bork^{1,3}

DNA microarray studies have shown that hundreds of genes are transcribed periodically during the mitotic cell cycle of humans¹, budding yeast^{2,3}, fission yeast^{4–6} and the plant *Arabidopsis thaliana*⁷. Here we show that despite the fact the protein complexes involved in this process are largely the same among all eukaryotes, their regulation has evolved considerably. Our comparative analysis of several large-scale data sets reveals that although the regulated subunits of each protein complex are expressed just before its time of action, the identity of the periodically expressed proteins differs significantly between organisms. Moreover, we show that these changes in transcriptional regulation have co-evolved with post-translational control independently in several lineages; loss or gain of cell-cycle-regulated transcription of specific genes is often mirrored by changes in phosphorylation of the proteins that they encode. Our results indicate that many different solutions have evolved for assembling the same molecular machines at the right time during the cell cycle, involving both transcriptional and post-translational layers that jointly control the dynamics of biological systems.

To obtain comparable sets of transcriptionally regulated genes from distantly related eukaryotes, we reanalysed the existing cell-cycle gene expression data^{1–7} and found 600 periodically expressed genes in *Homo sapiens*, 600 in *Saccharomyces cerevisiae*^{8,9}, 500 in *Schizosaccharomyces pombe*¹⁰ and 400 in *Arabidopsis thaliana*. These gene lists are more conservative than those originally proposed, but achieve better sensitivity (estimated to be 80–90% for human and the two yeasts, but only 50% in the plant; see Supplementary Information). We further assigned genes with a common descent to orthologous groups, of which we estimate at least 80% to contain the correct set of functionally equivalent genes (see Supplementary Information). The complete list of orthologous groups with periodic members and their peak times is available at <http://www.cbs.dtu.dk/cellcycle>.

Of these groups, 381 contain orthologues from all four organisms and have at least one periodic member. This conserved set includes 550 of the total set of 2,100 periodically expressed genes, implying that most of the periodically expressed genes do not have orthologues in all four organisms. Even among this subset we found that periodicity is poorly conserved across the four organisms (Fig. 1), meaning that although the protein sequences are conserved through evolution, their transcriptional regulation during the cell cycle is not. This is compatible with the results of earlier comparisons of periodically expressed genes from two or three organisms^{4,5,10–13}. Although the true overlap might be slightly better than is shown in Fig. 1, the large differences cannot be explained by the quality of the gene expression data or the orthology analysis (see Supplementary Information).

Only five orthologous groups are periodic in all four organism, namely a group of mitotic cyclins, three groups of histone proteins, and subunits of the ribonucleotide-diphosphate reductase (RNR) complex. The three additional orthologous groups that are periodic

in all organisms except *A. thaliana* consist of the CDC20 recognition subunit of the anaphase-promoting complex/cyclosome (APC/C), the pre-replication complex component CDC6, and the sister chromatid cohesion protein MCD1. The genes for which periodicity is highly conserved thus encode either key regulators of the cell cycle or components that are needed for synthesizing the building blocks of new DNA and chromatin.

For the genes for which periodic expression *per se* is conserved, we compared whether their expression peaks during the same phase of the cell cycle in each organism (Fig. 1). Despite normalizations to correct for organism-specific differences in the relative lengths of the cell-cycle phases (see Supplementary Information), we found that the timing of expression is only conserved for 42% of the orthologous groups (allowing a relaxed 20% tolerance, Fig. 1). The widespread differences in both identity and timing of the transcriptionally

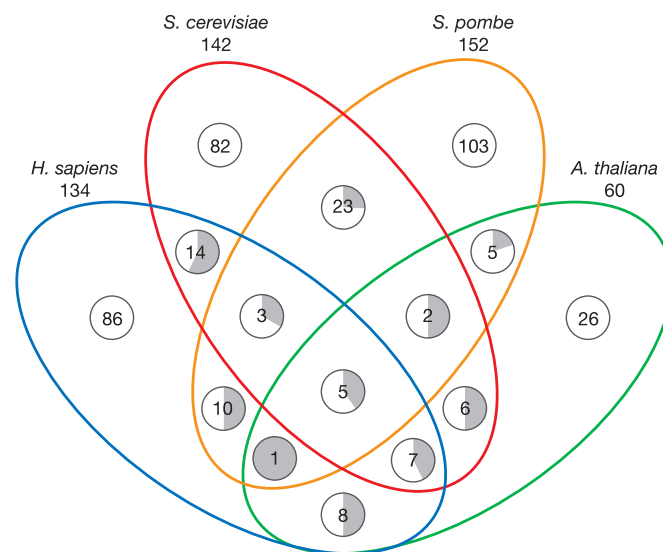


Figure 1 | Periodic expression is poorly conserved. The four-way Venn diagram summarizes the conservation of periodic expression for the 381 orthologous groups that contain at least one protein from each of the four organisms, and where one or more members are periodically expressed. The total number of such orthologous groups for each organism is given below its name. The pie charts show the fraction of orthologous groups for which the peak time is conserved (grey) or has changed by more than one-fifth of a cell cycle (white). Together, these analyses reveal that periodic expression is poorly conserved at the level of individual genes: conserved periodic expression across the four organisms is observed in only five cases and for only two of these is the timing conserved as well, namely histones H2A and H4 (see also Supplementary Figs 6 and 7).

¹European Molecular Biology Laboratory, D-69117 Heidelberg, Germany. ²Center for Biological Sequence Analysis, Technical University of Denmark, DK-2800 Lyngby, Denmark.

³Max-Delbrück-Centre for Molecular Medicine, D-13092 Berlin, Germany.

*These authors contributed equally to this work.

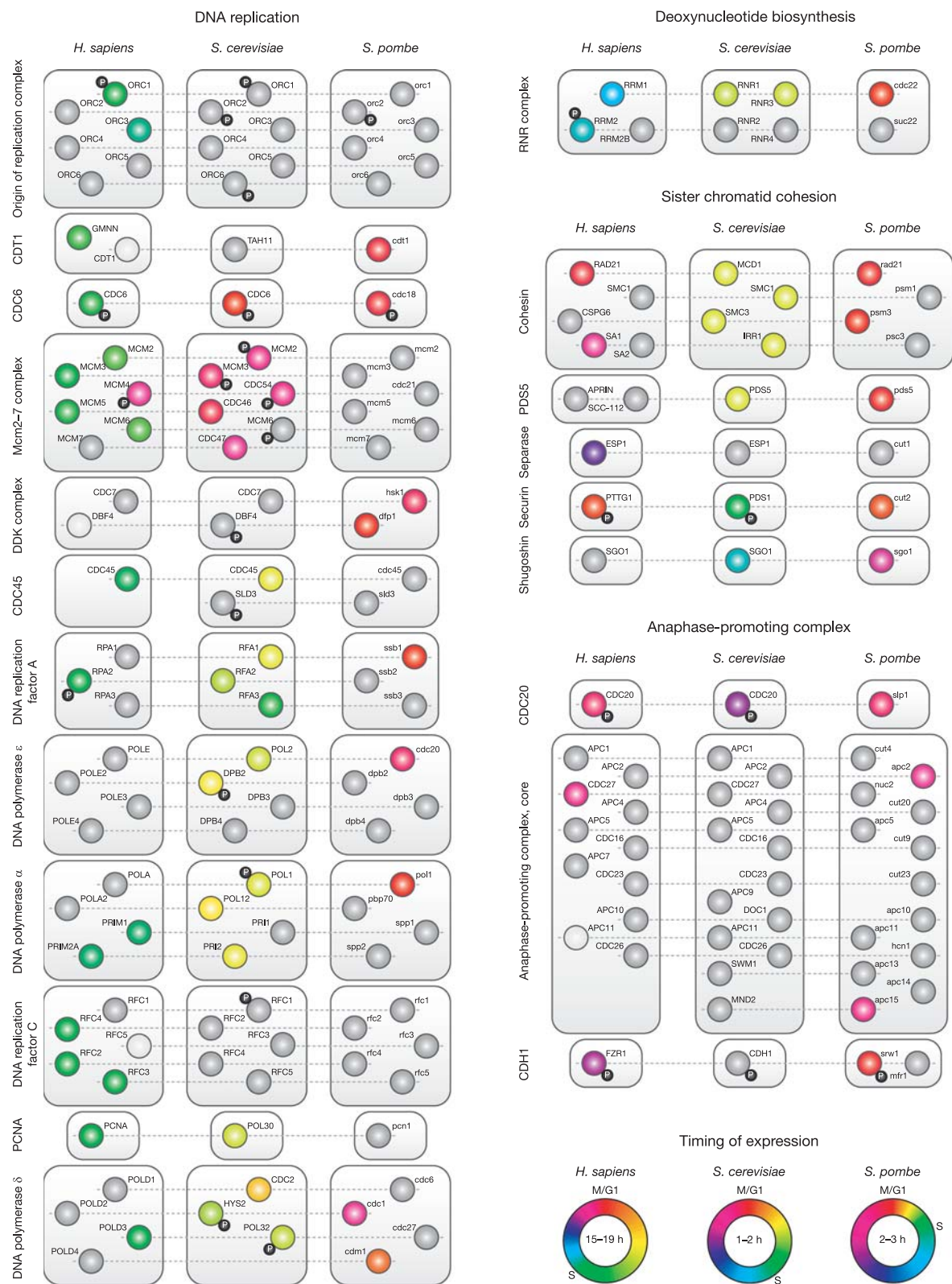


Figure 2 | Protein complexes are regulated through different subunits in each organism. The best-studied protein complexes involved in core cell-cycle processes were curated from the literature and public databases (see Supplementary Information). The subunits (circles) of each protein complex (boxes) are shown for three organisms (columns). Proteins on the same horizontal line are sequence orthologues and are believed to perform the same function. Dynamic subunits are coloured according to their time of peak expression, static subunits are shown in grey, and subunits for which

no expression data were available are displayed in white. As the relative length of the phases varies between organisms, time warping was used to ensure that the same colour corresponds to the same stage of the cell cycle (see colour legends). The small black circles denote proteins that experiments indicate are phosphorylated by CDKs. Detailed discussions on the composition and regulation of these and other complexes can be found in Supplementary Information, along with details on the time warping.

regulated genes are at first glance surprising, considering the high degree of conservation of the core cell-cycle machinery.

However, protein complexes that are involved in the cell cycle often comprise a mixture of static (constitutively expressed) and dynamic (periodically expressed) subunits, the latter being expressed just before the stage of the cell cycle in which the complex is active ('just-in-time assembly' of complexes)⁹. To confirm the consistency of the periodic expression of complex subunits within each organism, and to test a previously stated hypothesis that the just-in-time assembly principle allows the identity of the regulated subunits to change during evolution⁹, we carefully annotated the composition of some of the best-understood cell-cycle complexes in human and the two yeasts (Fig. 2; *A. thaliana* was omitted owing to the poor sensitivity of the expression data and the lack of knowledge on complex composition).

When mapping periodically expressed genes onto equivalent protein complexes, it becomes obvious that the dynamic subunits of individual complexes are consistently co-expressed, and that the timing of expression is similar for complexes that are involved in the same cellular process within each organism but differs between organisms (Fig. 2). However, our analysis also shows that the identity of the dynamic subunits within each complex has changed during evolution. The differences in temporal expression of individual, orthologous genes can thus be understood when viewed at the level of protein complexes and cellular processes (Fig. 2).

Furthermore, the differences in temporal expression are in remarkable agreement with current knowledge about the order and timing of assembly of the individual complexes that are involved in DNA replication^{14,15} and mitosis¹⁶, and even provides new insight into the regulation of these complexes. For example, both the sister chromatid cohesion complex and securin, which prevents separate from cleaving cohesin, are expressed during mitosis in human and fission yeast but at the G1/S transition in budding yeast. The same evolutionary changes in timing are observed for Pds5 and shugoshin, supporting the recent suggestions that Pds5 is a fifth subunit of cohesin and that shugoshin might help to protect mitotic cohesin from degradation (see Supplementary Information)¹⁶. The regulation of the DNA replication machinery has also evolved considerably. In fission yeast, it is expressed as early as M phase (Fig. 2) owing to the very

short G1 phase^{17,18}. Also, only two subunits of the pre-replication complex are subject to transcriptional regulation, namely *cdt1* and *cdc18*, which when overexpressed can induce re-replication^{18,19}. In humans, the pre-replication complex is regulated through many more subunits. Surprisingly, MCM4 is expressed in M phase, whereas the rest of the dynamic subunits appear in G1/S, which has to our knowledge never been described before. The distinct transcriptional regulation of MCM4 is intriguing, considering that phosphorylation of this subunit regulates the chromatin association of the entire Mcm2–7 complex^{20,21}. Many more details on individual complexes are presented in Supplementary Information.

The agreement between the timing of transcription and the timing of action of the protein products is not necessarily to be expected, as phosphorylation is at least as important as transcription for regulating cell-cycle complexes. In this regard, we have noted that periodically expressed *S. cerevisiae* proteins tend also to be phosphorylated⁹. To generalize this observation, we annotated the complexes in Fig. 2 with information on phosphorylation by cyclin-dependent kinases, and found that the dynamic subunits are three times as likely to be targeted by phosphorylation as the static ones ($P < 10^{-3}$).

To assess the evolutionary significance of this correlation, we applied two proteome-wide statistical tests to several independent sets of known and predicted phosphoproteins from each organism (Fig. 3). We first compared the phosphorylation of dynamic and static proteins from each organism (Fig. 3a). This comparison revealed highly significant over-representation of dynamic proteins among human cyclin-dependent kinase (CDK) substrates identified by small-scale experiments²², among human phosphoproteins identified by mass-spectrometry studies, and among *S. cerevisiae* Cdc28 substrates identified in at least one of two systematic screens^{23,24}. Sequence-based predictions²⁵ provide yet another complementary, unbiased source of CDK phosphorylation sites, which supports the experimental results and shows that the correlation holds true also in fission yeast (Fig. 3a).

The fact that CDK substrates are enriched among the dynamic proteins in three organisms could be due to a core set of dynamic orthologues that are phosphorylated by CDKs in all organisms or, alternatively, it could reflect that loss or gain of transcriptional

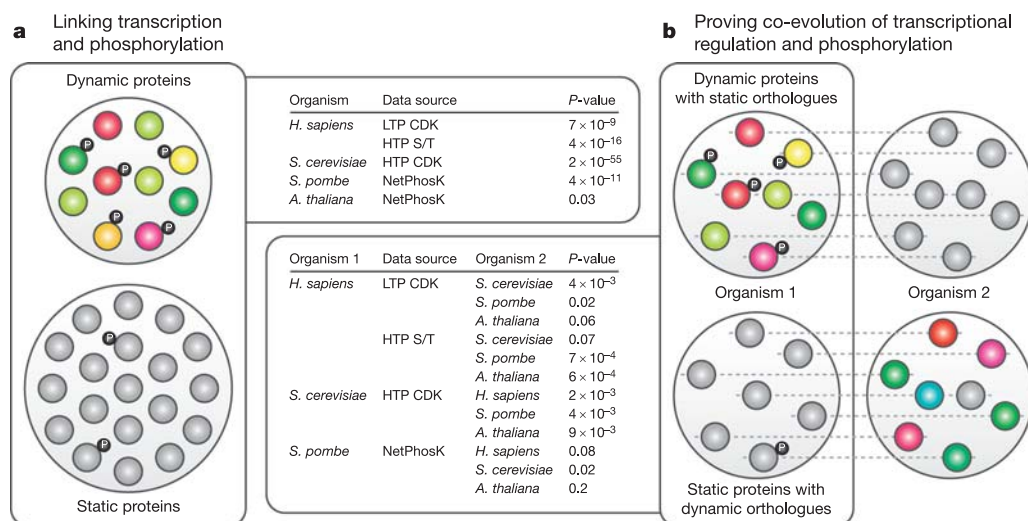


Figure 3 | Cyclin-dependent kinases preferentially phosphorylate dynamic proteins. **a, b**, Two statistical tests were used to show that transcriptional and post-translational regulation tend to control the same proteins. The test shown in **a** compares the dynamic proteins in an organism to all other (static) proteins encoded by the genome. The much stricter test shown in **b** instead compares only the dynamic proteins with static orthologues to the static proteins with dynamic orthologues, thereby directly showing that loss or gain of transcriptional regulation correlates with loss or gain of phosphorylation. Fisher's exact test was used to calculate the statistical significance of the

over-representation among the dynamic proteins of phosphoproteins that are either known CDK substrates from low-throughput experiments (LTP CDK²²), contain S/T phosphorylation sites according to high-throughput (HTP) mass-spectrometry studies (HTP S/T²²), were identified as CDK substrates in systematic screens (HTP CDK^{23,24}), or were predicted to contain CDK phosphorylation sites on the basis of their sequences (NetPhosK²⁵). The lower significance in the second, stricter test mainly reflects lower counts, rather than changes in the fractions of phosphoproteins among the dynamic versus the static proteins (see Supplementary Information).

regulation of a gene is correlated with loss or gain of phosphorylation of the corresponding protein. Examples of the latter include the human RRM2 subunit of the RNR complex and the budding yeast DPB2 subunit of DNA polymerase- ϵ (Fig. 2), which are both dynamic and phosphorylated whereas their orthologues are neither (see Supplementary Information). Within each organism, we therefore compared the dynamic proteins with static orthologues to static proteins with dynamic orthologues (Fig. 3b) and found that transcriptional regulation and phosphorylation have co-evolved in the three organisms studied. Notably, this co-evolution is supported by all pairwise organism comparisons and by both experimental data and sequence-based predictions. Dynamic proteins are thus often controlled through both mechanisms (Fig. 2), although the identity of the dynamic proteins varies greatly between organisms (Fig. 1).

We have previously found in *S. cerevisiae*⁹ that, like CDK phosphorylation, cell-cycle-regulated proteolysis preferentially affects the dynamic proteins. Our current results raise the intriguing possibility that all three levels of regulation have co-evolved. Unfortunately, experimental data on cell-cycle-regulated proteolysis is scarce, and we therefore applied the two statistical tests described above to sets of proteins whose sequences contain putative degradation signals. Although these data are too weak to prove co-evolution, the results show that targeted degradation preferentially affects the transcriptionally regulated proteins in all four organisms (see Supplementary Information).

It is tempting to speculate on the driving force that leads to the co-evolution of transcriptional and post-translational regulation, which we could prove in three of the four separate lineages studied. Requiring both transcription and phosphorylation of the same key components might increase robustness to prevent accidental activation. For other proteins, transcriptional regulation might be sufficient for activation, whereas phosphorylation might cause inactivation, for example, by targeting them for degradation.

Together, our results provide a first global view of the evolutionary dynamics of the transcriptional and post-translational regulation of a large and complex biological system. They clearly indicate that although the same general underlying principles, namely just-in-time assembly and multi-layer regulation of functional modules, are widely conserved in eukaryotes, the detailed regulation of individual genes and proteins varies greatly and thus generally cannot be inferred from distantly related organisms. This raises the question of how fast regulation evolves and how closely related two organisms have to be for regulatory details to be transferable. Although microarray expression studies from more closely related model organisms will be needed to investigate this globally, we show that changes in regulation can take place in the order of only a hundred million years (see examples in Supplementary Information), implying that even within vertebrates regulation might differ considerably.

METHODS

Analysis of microarray expression data. Periodically expressed genes were identified by applying the same computational method⁸ to 19 cell-cycle microarray timecourses^{1–7}, and the results were benchmarked against genes for which other evidence indicates periodic expression. For each periodic gene, the time of peak expression was calculated, and the time scales were made comparable across organisms through time warping (see Supplementary Information).

Identification of orthologous proteins. All-against-all Smith–Waterman similarities were calculated, and the proteins were grouped into orthologous groups using a triangular linkage clustering similar to the original COG (Clusters of Orthologous Groups) procedure²⁶ (see Supplementary Information). The quality of these orthologous groups was assessed on the basis of the manually assigned orthology relationships for DNA replication complexes in Fig. 2.

Curation of complexes. The composition of the individual complexes in Fig. 2 was based on literature and annotations from the relevant model organisms databases (SGD for budding yeast; GeneDB for fission yeast; Ensembl, Uniprot and Reactome for human). Details on the annotations can be found in Supplementary Information and at <http://www.cbs.dtu.dk/cellcycle>.

Correlation analyses of transcriptional regulation and phosphorylation. Lists

of human and budding yeast CDK phosphoproteins were compiled from the Phospho.ELM database²² and large-scale screens^{23,24}. In addition, CDK substrates were predicted from protein sequences using the NetPhosK cdk5 method²⁵, which we benchmarked on known CDK substrates (see Supplementary Information). A set of human phosphoproteins (not CDK-specific) was compiled from high-throughput mass-spectrometry studies²². For each of these sets of phosphoproteins, Fisher's exact test was used to assess the statistical significance of the correlation between transcriptional regulation and phosphorylation.

Received 23 May; accepted 18 August 2006.

Published online 27 September 2006.

- Whitfield, M. L. *et al.* Identification of genes periodically expressed in the human cell cycle and their expression in tumors. *Mol. Biol. Cell* **13**, 1977–2000 (2002).
- Cho, R. J. *et al.* A genome-wide transcriptional analysis of the mitotic cell cycle. *Mol. Cell* **2**, 65–73 (1998).
- Spellman, P. T. *et al.* Comprehensive identification of cell cycle-regulated genes of the yeast *S. cerevisiae* by microarray hybridization. *Mol. Biol. Cell* **9**, 3273–3297 (1998).
- Rustici, G. *et al.* Periodic gene expression program of the fission yeast cell cycle. *Nature Genet.* **36**, 809–817 (2004).
- Peng, X. *et al.* Identification of cell cycle-regulated genes in fission yeast. *Mol. Biol. Cell* **16**, 1026–1042 (2005).
- Oliva, A. *et al.* The cell cycle-regulated genes of *Schizosaccharomyces pombe*. *PLoS Biol.* **3**, e225 (2005).
- Menges, M., Hennig, L., Grussem, W. & Murray, J. A. H. Genome-wide gene expression in an *Arabidopsis* cell suspension. *Plant Mol. Biol.* **53**, 423–442 (2003).
- de Lichtenberg, U. *et al.* Comparison of computational methods for the identification of cell cycle regulated genes. *Bioinformatics* **21**, 1164–1171 (2005).
- de Lichtenberg, U., Jensen, L. J., Brunak, S. & Bork, P. Dynamic complex formation during the yeast cell cycle. *Science* **307**, 724–727 (2005).
- Marguerat, S. *et al.* The more the merrier: comparative analysis of microarray studies on cell cycle-regulated genes in fission yeast. *Yeast* **23**, 261–277 (2006).
- Ota, K., Goto, S. & Kanehisa, M. Comparative analysis of transcriptional regulation in eukaryotic cell cycles. *Proc. Fourth Intl Workshop Bioinformatics Syst. Biol.* 26–27 (2004).
- Sherlock, G. STARTing to recycle. *Nature Genet.* **36**, 795–796 (2004).
- Dyczkowski, J. & Vingron, M. Comparative analysis of cell cycle regulated genes in eukaryotes. *Genome Inform. Ser. Workshop Genome Inform.* **16**, 125–131 (2005).
- Bell, S. P. & Dutta, A. DNA replication in eukaryotic cells. *Annu. Rev. Biochem.* **71**, 333–374 (2002).
- Kearsey, S. E. & Cotterill, S. Enigmatic variations: divergent modes of regulating eukaryotic DNA replication. *Mol. Cell* **12**, 1067–1075 (2003).
- Nasmyth, K. & Haering, C. H. The structure and function of SMC and kleisin complexes. *Annu. Rev. Biochem.* **74**, 595–648 (2005).
- Forsburg, S. L. & Nurse, P. Cell cycle regulation in the yeasts *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*. *Annu. Rev. Cell Biol.* **7**, 227–256 (1991).
- Diffley, J. F. X. Regulation of early events in chromosome replication. *Curr. Biol.* **14**, R778–R786 (2004).
- Yanow, S. K., Lygerou, Z. & Nurse, P. Expression of Cdc18/Cdc6 and Cdt1 during G2 phase induces initiation of DNA replication. *EMBO J.* **20**, 4648–4656 (2001).
- Ekholm-Reed, S. *et al.* Deregulation of cyclin E in human cells interferes with prereplication complex assembly. *J. Cell Biol.* **165**, 789–800 (2004).
- Zhu, Y., Ishimi, Y., Tanudji, M. & Lees, E. Human CDK2 inhibition modifies the dynamics of chromatin-bound minichromosome maintenance complex and replication protein A. *Cell Cycle* **4**, 1254–1263 (2005).
- Diella, F. *et al.* Phospho.ELM: a database of experimentally verified phosphorylation sites in eukaryotic proteins. *BMC Bioinformatics* **5**, 79 (2004).
- Ubersax, J. A. *et al.* Targets of the cyclin-dependent kinase cdk1. *Nature* **425**, 859–864 (2003).
- Loog, M. & Morgan, D. O. Cyclin specificity in the phosphorylation of cyclin-dependent kinase substrates. *Nature* **434**, 104–108 (2005).
- Blom, N., Sicheritz-Ponten, T., Gupta, R., Gammeltoft, S. & Brunak, S. Prediction of post-translational glycosylation and phosphorylation of proteins from the amino acid sequence. *Proteomics* **4**, 1633–1649 (2004).
- Tatusov, R. L., Koonin, E. V. & Lipman, D. J. A genomic perspective on protein families. *Science* **278**, 631–637 (1997).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank C. von Mering for assistance with detection of orthologues, and P. Nurse, E. Karsenti, J. Bähler, and members of the Bork and Brunak groups for comments on the manuscript. This work was supported by grants from the Danish National Research Foundation, the Danish Technical Research Council, and the European Commission FP6 Programme (grants DIAMONDS and BioSapiens).

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to P.B. (bork@embl.de).

ERRATUM

doi:10.1038/nature05187

Coral reef diversity refutes the neutral theory of biodiversity

Maria Dornelas, Sean R. Connolly & Terence P. Hughes

Nature 440, 80–82 (2006)

The original data from this paper are now available online as additional Supplementary Information to the original paper. They should therefore be cited as part of the original paper. The coral abundance distributions are shown in Supplementary Table 1. The similarity matrix is shown as Supplementary Table 2.

Supplementary Information is linked to the online version of the original paper at www.nature.com/nature.

CORRIGENDUM

doi:10.1038/nature05189

Sexual reproduction selects for robustness and negative epistasis in artificial gene networks

Ricardo B. R. Azevedo, Rolf Lohaus, Suraj Srinivasan, Kristen K. Dang & Christina L. Burch

Nature 440, 87–90 (2006).

Equation (1) in this paper was incorrect. The correct equation is:

$$s_i(t+1) = f \left[\sum_{j=1}^N r_{ij} s_j(t) \right] \quad (1)$$

The equation given in Section 1.5 of the Supplementary Information (page 4) was also incorrect (it lacked the minus sign). The Supplementary Information has now been updated. All results reported in the paper were obtained using the correct equations.

Supplementary Information is linked to the online version of the original paper at www.nature.com/nature.

CORRIGENDUM

doi:10.1038/nature05211

Arctic hydrology during global warming at the Palaeocene/Eocene thermal maximum

Mark Pagani, Nikolai Pedentchouk, Matthew Huber, Appy Sluijs, Stefan Schouten, Henk Brinkhuis, Jaap S. Sinninghe Damsté, Gerald R. Dickens & the Expedition 302 Scientists (Jan Backman, Steve Clemens, Thomas Cronin, Frédérique Eynaud, Jérôme Gattacceca, Martin Jakobsson, Ric Jordan, Michael Kaminski, John King, Nalân Koc, Nahysa C. Martinez, Jens Matthiessen¹, David McInroy, Theodore C. Moore Jr, Kathryn Moran², Matthew O'Regan, Jonaotaro Onodera, Heiko Pälike, Brice Rea, Domenico Rio, Tatsuhiko Sakamoto, David C. Smith, Ruediger Stein¹, Kristen E. K. St John, Itsuki Suto, Noritoshi Suzuki, Kozo Takahashi, Mahito Watanabe & Masanobu Yamamoto)

¹Department of Geosystems, Alfred Wegener Institute for Polar and Marine Research (AWI), PO Box 120161, D-27515 Bremerhaven, Germany. ²Graduate School of Oceanography and Department of Ocean Engineering, University of Rhode Island, Narragansett Bay Campus, Narragansett, Rhode Island 02882, USA.

Nature 442, 671–675 (2006)

We omitted the names of the following authors of this Letter: Jens Matthiessen, Kathryn Moran and Ruediger Stein.

CORRIGENDUM

doi:10.1038/nature05153

Microstimulation of inferotemporal cortex influences face categorization

Seyed-Reza Afraz, Roozbeh Kiani & Hossein Esteky

Nature 442, 692–695 (2006)

In Fig. 1a the number in the right lower section in front of the term 'visual signal' below the last face image should be 80% and not 20%.

naturejobs

The announcement of the 2006 Nobel laureates, which started this week, is the culmination of the science award season. The Nobel, the Lasker and other prestigious prizes have all been recently awarded and all honour achievements in science.

But there is a newer crop of less august, yet increasingly important, awards that value potential as much as past work. They highlight people who change the way science is done, and can be won by scientists at many different stages in their careers.

The European Union Contest for Young Scientists rewards undergraduate initiative and encourages young people to consider science as a viable profession. This year there are three winners of the €5,000 (US\$6,000) prize. One group created a system to de-ice aeroplanes; the second discovered a way to study the flight curves of table-tennis balls; and the third synthesized a potential beta blocker.

Winners of Nobels and Laskers are free to pursue blue-sky, hypothesis-driven research. But few younger, mid-career scientists have that option. The US National Institutes of Health's Pioneer Award changes that, by giving its recipients US\$2.5 million over five years to pursue unconventional ideas that might not fare well in the traditional peer-review process.

Finally, mentors who help younger scientists to launch careers often don't get the honours they would have got if they had focused on their own. The UK National Endowment for Science, Technology and the Arts mentoring prize, co-sponsored by *Nature*, tries to redress the balance. It is known as the 'unsung hero' award and was created to highlight the methods of both senior and mid-career researchers who work hard on advancing their trainees' careers as much as their own.

Although scientists will always dream of Nobels, fantasizing about these other awards and creating similar ones could improve science for themselves and for future generations.

Paul Smaglik, *Naturejobs* editor

Publisher: Ben Crowe
Editor: Paul Smaglik
Assistant Editor: Gene Russo

European Head Office, London
The Macmillan Building,
4 Crinan Street, London N1 9XW, UK
Tel: +44 (0) 20 7843 4961
Fax: +44 (0) 20 7843 4996
e-mail: naturejobs@nature.com

European Sales Manager:
Andy Douglas (4975)
e-mail: a.douglas@nature.com
Business Development Manager:
Amelie Pequignot (4974)
e-mail: a.pequignot@nature.com

Natureevents:
Claudia Paulsen Young (+44 (0) 20 7014 4015)
e-mail: c.paulsenyoung@nature.com
France/Switzerland/Belgium:
Muriel Lestringuez (4994)

UK/Ireland/Italy/RoW:
Loredana Milanese (4944)
Nils Moeller (4953)
Scandinavia/Spain/Portugal:
Evelina Rubio-Morgan (4973)
Germany/Austria/The Netherlands:
Reya Silao (4970)
Online Job Postings:
Matthew Ward (+44 (0) 20 7014 4059)

European Satellite Office
Germany: Patrick Phelan
Tel: +49 89 54 90 57 11
Fax: +49 89 54 90 57 20
e-mail: p.phelan@nature.com

Advertising Production Manager:
Stephen Russell
To send materials use London address above.
Tel: +44 (0) 20 7843 4816
Fax: +44 (0) 20 7843 4996
e-mail: naturejobs@nature.com

Naturejobs web development: Tom Hancock
Naturejobs online production:
Catherine Alexander

US Head Office, New York
75 Varick Street, 9th Floor,
New York, NY 10013-1917
Tel: +1 800 989 7718
Fax: +1 800 989 7103
e-mail: naturejobs@natureny.com

US Sales Manager: Peter Bless

Japan Head Office, Tokyo
Chiyoda Building, 2-37 Ichigayatamachi,
Shinjuku-ku, Tokyo 162-0843
Tel: +81 3 3267 8751
Fax: +81 3 3267 8746

Asia-Pacific Sales Manager:
Ayako Watanabe
e-mail: a.watanabe@natureasia.com

**THE CAREERS
MAGAZINE FOR
SCIENTISTS**

A jumping off point



The field of evolutionary genetics is itself in the midst of astonishingly rapid evolution. For decades limited largely to theoretical considerations and a gene-by-gene focus, the field blossomed after the first genome was sequenced in 1995. Since then, researchers have been able to pose new types of questions and pursue the answers with experiments barely imaginable a generation ago.

Armed with hundreds of sequenced genomes, today's evolutionary geneticists tackle problems as diverse as tracing the origins of infectious diseases, explaining rampant obesity and conserving endangered species. The field actually embraces two research areas: how genes underlie evolutionary processes, and how genes and genomes themselves evolve.

Such broad questions invite comparisons across species and data sets, says Kateryna Makova, a biologist at the large Center for Comparative Genomics and Bioinformatics at Pennsylvania State University in University Park. Even a small department will offer varied research opportunities.

The five-member School of Population and Evolutionary Genetics at the University of Nottingham, UK, for example, investigates bacteria, insects, snails, slugs and a 'frozen ark' of stored DNA from thousands of endangered animal species.

Modern evolutionary biology is built on a foundation of genetics. Powerful new tools complement the imaginations of researchers intent on teasing out clues to the past from genes and genomes. And there is

The recent flood of genome sequences has given evolutionary genetics a boost.

Ricki Lewis takes a sharp look at a varied field.

no shortage of data to mine. As of September 2006, a total of 429 complete genome sequences have been published, according to the Genomes OnLine Database, and ongoing projects will add another 1,000 (see 'The role of genomics'). "It will be very interesting to adapt the longstanding theoretical understanding of the interactions among selection, migration, genetic drift, mutation and demographic change to deal with these very large data sets," says John Brookfield, an evolutionary geneticist at the University of Nottingham.

In the money

Funding is robust. In 2006 in the United States, for example, the National Human Genome Research Institute (NHGRI) budgeted \$133 million for its five-centre large-scale sequencing research network, according to agency spokesman Geoff Spencer. The spectacular growth of The Institute for Genomic Research (TIGR) in Rockville, Maryland, launched in 1992, illustrates just how dramatically the field is booming. A not-for-profit research organization funded by private and government sources and home of the first genome sequence — that of *Haemophilus influenzae* — TIGR's Comprehensive Microbial Database now boasts more than 300 sequenced genomes.

The time element distinguishes evolutionary genetics from much of biology. Most geneticists are concerned with how present-day adaptations protect individuals. An evolutionary geneticist, explains Brookfield, looks backwards, hypothesizing how modern gene or genome structure reflects selection driven by past environmental factors. "Evolutionary genetics gives us a 'common yardstick', a scale that applies to all measurements of genetic change," says David Lunt, a molecular ecologist at the University of Hull, UK.

Makova says that evolutionary geneticists need the "ability to integrate complex concepts from

From the lynx to the gibbon: today's study of genes and evolution is wide-ranging.



J. ZUCKERMAN, GIBBON CONSERVATION CENTER

THE ROLE OF GENOMICS

Today's young evolutionary biologists may not remember a time when genome sequences weren't readily available. However, having such sequences at one's fingertips is not an endpoint in the research process, but a new type of beginning. Creativity and imagination are still key to success in evolutionary genetics.

Eventually the correlative data obtained through evolutionary bioinformatics will have to be rigorously tested in experimental models, notes Petr Tvrdik, a postdoctoral fellow at the University of Utah. Still, he emphasizes that genome sequences are critical in filling out the branches of evolutionary trees.

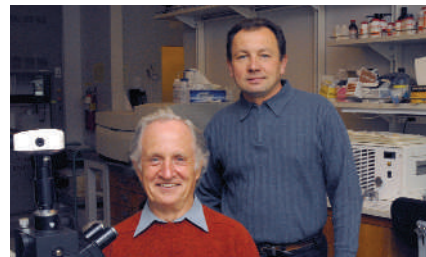
Angus Davison of the University of Nottingham has designed several courses in

evolutionary genetics; he has found expressed sequence tags (ESTs), which herald protein-encoding genes, are more useful than entire sequenced genomes. He works on molluscs, not one of which has so far had its genome sequenced. "It is not just the sequenced genomes that are providing the impetus — the sheer volume of sequences, especially ESTs, from non-model organisms has had a big impact," he says.

Although he appreciates what genomics has done for biology, evolutionary biologist David Lunt challenges its value in evolutionary research. "Although genomic projects produce a truly impressive quantity of data, they have not been designed to address evolutionary

questions," says Lunt, based at the University of Hull, UK. "I think we have got wrapped up in sound bites and buzzwords and vague possibilities for the future without really seeing what it can actually do for us now."

R.L.



Petr Tvrdik (right) and his lab chief Mario Capecchi.

different disciplines, such as biology, statistics and computer science". And just as a species survives because of its ability to adapt, a successful evolutionary geneticist must adapt to a "diversity of software and computer platforms", says Blair Hedges, a biologist at Pennsylvania State University. Other skills are vital too. "It takes persistence. I think the most important skill is to read a lot, understand and remember," says Lunt.

Research bridging evolution and genetics may take place at a computer or lab bench, or in the field. A project at the University of Utah School of Medicine in Salt Lake City illustrates the creativity that can drive experimental research. Postdoctoral fellow Petr Tvrdik has reversed what may have happened over evolutionary time, by generating mice with a non-functioning homeobox (*hox*) gene that would normally control facial expression, then restoring it by adding the regulatory region of a related *hox* gene.

His experiments tests the long-held theory that members of gene families duplicated and diverged from a single gene over time, some members retaining the ability to affect others. "Our example shows that critical theoretical information is required in combination with experimental expertise to tackle the complex questions of evolutionary genetics," says Tvrdik.

Diverse directions

Research in evolutionary genetics takes many forms. For instance, any experiment that uses a gene in a model organism to discover a similar gene in a different species (such as our own) uses an evolutionary genetics approach. Genome similarities have driven the NHGRI to include diverse species in the genome sequencing projects it funds. The most recent candidate, a gibbon, will join five other non-human primates. Also slated for genome sequencing are ten modern species whose ancestors straddle the boundary between unicellularity and multicellularity.

Many aspects of human health and disease have arisen from the temporal tango engaged in by our genomes and the environment. Current epidemics of diabetes and obesity, for example, may reflect natural selection of gene variants that fostered nutrient storage during long-ago famines. Understanding viral diversity and evolution has helped identify the cause of severe acute respiratory syndrome (SARS). "If we can sequence a gene, put it into a tree of life, and find its close relatives, we can then begin to understand what type of beast our new disease is," says Lunt. Even Tvrdik's reconstruction

of a half-billion-year-old gene in mice may have an impact on human health care by suggesting therapy that coaxes one gene to rescue another.

Conservation biology also uses evolutionary genetics. At the University of Uppsala in Sweden, evolutionary biologist Hans Ellegren and his group reconstruct histories of endangered wolves and lynxes from DNA in the excrement and hair of modern animals and 200-year-old museum specimens. The Swedish Research Council has funded him to found a Centre of Excellence in Evolutionary Genomics.

And the discipline is aiding our understanding of cancer. Marianne Bronner-Fraser, who heads the Centre for Excellence in Genomic Science at the California Institute of Technology in Pasadena, is examining neural-crest formation in zebrafish and chickens — models for more complex vertebrates. Human diseases that stem from neural-crest anomalies include melanoma, neuroblastoma and neurofibromatosis. The funding for the centre is \$18 million over the next five years.

Marianne Bronner-Fraser and Angus Davison revel in the field of evolutionary genetics.



Branching out

In training for a career in evolutionary genetics, the department name is not as important as faculty interactions. "A narrow group, such as our 'population and evolutionary genetics' group, within a broad department, is probably best," says Angus Davison of the Institute of Genetics at the University of Nottingham. Brookfield suggests joining a genetics department that offers different subdisciplines, such as developmental genetics and palaeobiology. And Makova thinks that getting faculty members actively interacting and working in the area is key. "It can be in the same department, but it doesn't have to be. We have a centre with weekly meetings," says Makova, who draws from biology, biochemistry and molecular biology, anthropology, plant pathology, entomology and animal-science departments.

A career in evolutionary genetics weds the application of new tools and technologies to the imaginative exploration of how biodiversity came to be — an enticing goal. "The really important breakthroughs come from understanding a problem in a way nobody else does, and that is rarely accomplished by examining the minutiae more closely than others," says Lunt. "It's a holistic approach that is true of science in general."

Ricki Lewis is a freelance science writer in Albany, New York.



THE NATIONAL INSTITUTES OF HEALTH WWW.NIH.GOV

National Human Genome Research Institute Division of Intramural Research

From Base Pairs to Bedside

Although the completion of the Human Genome Project was a significant achievement, it was actually just the first step toward fulfilling the goal of improving human health through genetics-based studies. With this goal in mind, in 1993, the National Institutes of Health (NIH) established a dynamic, cutting-edge Intramural Program within the then-named National Center for Human Genome Research to serve as the focal point for genetics and genomics research at the NIH and worldwide. It was envisioned that this program would develop novel genomic expertise, technologies, and approaches that other research institutions, including other NIH Institutes, could then use in studying the various hereditary disorders afflicting humankind.

Today, the National Human Genome Research Institute (NHGRI) Division of Intramural Research is one of the premier research programs working to unravel the genetic basis of human disease. During its brief existence, this program has made many seminal contributions to the fields of genetics and genomics. Accomplishments of its investigators have included:

- Identification of the genes responsible for numerous human genetic diseases
- Development of new paradigms for mapping, sequencing, and characterizing the human and other vertebrate genomes
- Development and application of DNA microarray technologies for large-scale analyses of gene expression
- Creation of innovative computational tools for analyzing large quantities of genomic data
- Generation of animal models critical to the study of human inherited disorders
- Design of novel approaches for diagnosing and treating genetic diseases

Along with their collaborators, NHGRI investigators have embarked on a number of high-risk studies to unearth clues about the complex genetic pathways involved in human diseases. These efforts have used genomic sequence data from human and other species to pinpoint hundreds of potential disease genes, including those implicated in cancer, diabetes, premature aging, hereditary deafness, and various neurological, developmental, metabolic, and immunological disorders. These studies have brought together NHGRI basic scientists and clinicians in collaborations aimed at developing better approaches for detecting, diagnosing, and managing genetic diseases.

The NHGRI Division of Intramural Research plans and conducts a broad program of laboratory and clinical research on the main NIH campus in Bethesda, Maryland as well as at other sites in Maryland (such as the Bayview campus in Baltimore and the Twinbrook complex in Rockville). The Division is led by the NHGRI Scientific Director, with input from the Board of Scientific Counselors – an

external group that provides expert oversight for all ongoing research and training activities. The NHGRI Clinical Director provides oversight and support for all patient-based clinical research.

The major programs of the NHGRI Division of Intramural Research are organized around the following seven thematic Branches:

- Cancer Genetics Branch
- Genetic Disease Research Branch
- Genetics and Molecular Biology Branch
- Genome Technology Branch
- Inherited Disease Research Branch
- Medical Genetics Branch
- Social and Behavioral Research Branch

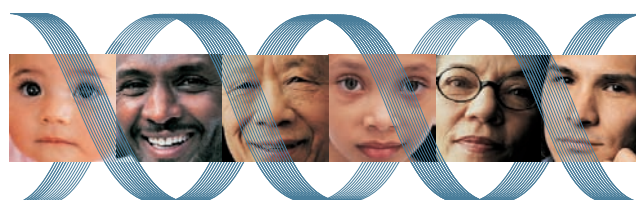
Each of the more than 40 NHGRI investigators is assigned to one of these Branches – although boundaries between Branches are artificial in many ways, with significant interactions among investigators and trainees occurring across the Division. There is also considerable overlap in the research areas emphasized by the different Branches. The NHGRI Division of Intramural Research is also supported by a number of outstanding scientific cores and centers.

NHGRI provides an excellent environment for training the next generation of researchers and clinicians, and is proud to support training at all career levels. The NHGRI Intramural Training Office serves as a focal point for training, offering a variety of resources and services, including career development seminars, grantsmanship training, matching trainees to individual research laboratories, and career and scientific mentoring.

Together, all of the elements of the NHGRI Division of Intramural Research share a common aim – to deliver on the promise of genetics and genomics by connecting the base pairs of the Human Genome Project to the bedside of those afflicted with genetic disease. We invite you to learn more about our program at <http://www.genome.gov/DIR>.

Eric D. Green, M.D., Ph.D., Scientific Director
Andy Baxevanis, Ph.D., Deputy Scientific Director

NATIONAL HUMAN GENOME RESEARCH INSTITUTE Division of Intramural Research





Department of Health and Human Services
National Institutes of Health
National Institute of Allergy and Infectious Diseases



With nation-wide responsibility for improving the health and well being of all Americans, the Department of Health and Human Services oversees the biomedical research programs of the National Institutes of Health (NIH) and those of NIH's research Institutes.

The National Institute of Allergy and Infectious Diseases (NIAID), a major research component of the NIH and the Department of Health and Human Services, is recruiting for a Tenure/Tenure Track position in the Laboratory of Host Defenses (LHD). The LHD studies immune functions essential for host defense against infection (inherited immune deficiencies) and those required for immune homeostasis (autoimmunity associated with excessive inflammation). The LHD seeks an M.D. or M.D., Ph.D. physician scientist to develop an independent translational research program related to the genetic basis, pathophysiology, diagnosis and treatment of autoimmune diseases associated with excessive inflammation. An emphasis on clinical aspects of innate immunity including phagocytic cells, natural killer cells, dendritic cells and other antigen presenting cells, toll-like receptors or other pattern recognition receptors in its interface with acquired immunity is desirable. The applicant should have a strong track record of basic research of the genetic basis of disease and alterations in signaling pathways responsible for immune dysregulation. The applicant must possess expertise and experience in the design and conduct of diagnostic and therapeutic clinical trials studying and treating autoimmune diseases. Strong clinical credentials in a specialty area relevant to the proposed translational research program (relevant specialties include but are not limited to rheumatology, pulmonary diseases, hematology, immunology or infectious diseases) are required. The program of study proposed by the applicant must include both laboratory components and the conduct of clinical protocols to assess new diagnostic and therapeutic modalities to diagnose and treat autoimmunity associated with excessive inflammation. Applicants particularly suitable for this program are those who have knowledge and experience in the development and clinical application of novel biological agents including chemokines, soluble chemokine receptors, adenosine receptor agonists, monoclonal antibodies, cellular therapies including transplantation or gene therapy to correct the abnormalities in immunity, that achieve immune tolerance or to reduce abnormal inflammation.

The applicant must provide evidence in the submitted materials that the applicant has a current license to practice medicine in one of the states of the United States or must have the all the credentials required by the State of Maryland for licensing to allow the practice of medicine. These credentials must include but are not limited to having a Doctor of Medicine or Doctor of Osteopathy degree from an accredited school in the U.S. or Canada, or a Doctor of Medicine or equivalent degree from a foreign medical school that provided education and medical knowledge substantially equivalent to accredited schools in the U.S. as demonstrated by permanent certification by the Educational Commission for Foreign Medical Graduates (ECFMG).

To be considered for this position, you will need to submit a curriculum vitae, bibliography, three (3) letters of reference, a detailed statement of research interests, and a hardcopy of selected publications to **Thomas A. Fleisher, MD, Chairperson, NIAID Search Committee, c/o Ms. Anissa N. Hunter, DIR Committee Coordinator, Reference Ad #009, 10 Center Drive MSC 1356, Building 10, Rm. 4A26, Bethesda, Maryland 20892-1356**. Completed applications **MUST** be received by **Thursday, November 15, 2006**. For additional information on this position, and for instructions on submitting your application, please see our website at: **www.niaid.nih.gov**.



THE NATIONAL INSTITUTES OF HEALTH

OPPORTUNITIES @ NIH



THE NATIONAL INSTITUTES OF HEALTH

OPPORTUNITIES @ NIH



**Department of Health and Human Services
National Institutes of Health
National Cancer Institute
Tenured, Tenure-Track Investigator**

With nationwide responsibility for improving the health and well being of all Americans, the Department of Health and Human Services (DHHS) oversees the biomedical research programs of the National Institutes of Health (NIH) and its research Institutes.

The National Cancer Institute (NCI), a major component of the NIH, is recruiting a tenure-track or tenured investigator to conduct primary research focused on human cancer risk related to infection or to perturbation of immunity or inflammation. The successful candidate will work in the Viral Epidemiology Branch, Division of Cancer Epidemiology and Genetics, NCI. Typical research projects are national or international in scope and involve collaborations with statisticians, clinicians, virologists, immunologists, geneticists, and other scientists throughout NIH and extramurally.

Candidates must have an M.D., Ph.D., or equivalent doctoral degree. They must have substantial experience in human viral, bacterial or cancer epidemiology. Demonstrated experience with the design, conduct, and analysis of epidemiologic studies is required. Candidates must have a sound working knowledge of basic statistical methods. Demonstrated experience with analyses of large databases, genomics, or genetic polymorphisms would be advantageous, as would clinical or laboratory skills. Candidates must have displayed strong leadership in independent, innovative research. They also should have outstanding communication skills for writing research papers, presenting at scientific meetings, and collaborating with scientists of various backgrounds. A substantial record of relevant publications in the peer-reviewed medical or scientific literature is required.

Applicants should send a cover letter that briefly summarizes their research interests and experience and future goals; curriculum vitae with bibliography; copies of three selected publications; and three letters of reference to: **Chair, Search Committee, Viral Epidemiology Branch, NCI, c/o Mrs. Judy Schwadron, Division of Cancer Epidemiology and Genetics, National Cancer Institute, Executive Plaza South, Room 8073, Bethesda, MD 20892** Or via email to: jrusell@mail.nih.gov

Candidates should submit applications by **November 15, 2006**, however, the search will continue until a qualified applicant is found.

**DEPARTMENT OF HEALTH AND HUMAN SERVICES (DHHS)
NATIONAL INSTITUTES OF HEALTH (NIH)
NATIONAL CANCER INSTITUTE (NCI)**



**HIV and AIDS Malignancy Branch
Center for Cancer Research**

**Tenured/Tenure Track Position
Translational Researcher in Viral Oncogenesis**

The HIV and AIDS Malignancy Branch (HAMB), NCI, is searching for a tenure-track or tenured investigator in the field of viral oncogenesis. It is anticipated that the investigator will establish an independent research program targeted to the study of the treatment, pathogenesis, and/or prevention of viral-induced tumors, especially those associated with AIDS. The research program should be translational in focus and be able to interface with a strong existing clinical research program in AIDS-related tumors. HAMB is located on the Bethesda campus of the NIH (<http://ccr.cancer.gov/labs/lab.asp?labid=63>). Current areas of research in HAMB focus on Kaposi's sarcoma-associated herpesvirus (KSHV/HHV-8)-associated tumors, the molecular biology of human papillomavirus (HPV), and the development of novel therapeutic interventions for HIV infection. Candidates for the position should have an M.D./Ph.D., Ph.D., or M.D. and strong research credentials. Applicants for this position should submit a curriculum vitae including bibliography, a statement of research interests, a two-page outline of the proposed research program, and the names of three references to **Chairman, Search Committee, HAMB, NCI, Attention Jan Huque, Building 10, Rm.10S255, 10 Center Drive, M.S.C. 1868, Bethesda, MD 20892-1868** no later than **December 7, 2006**. You may also e-mail your application to: huquei@mail.nih.gov (Jan Huque, 301-435-4627).



**Department Of Health And Human Services
National Institutes of Health
National Institute of Neurological Disorders and Stroke**

**Tenure Track/Tenure Position
Neurovirology
Division of Intramural Research**

The Division of Intramural Research of the National Institute of Neurological Disorders and Stroke is recruiting an individual for a tenure track position in the area of neurovirology with special interests and expertise in viruses that target the brain. Recruitment for a senior individual for a tenured position would also be considered. The individual would direct an independent research program on the molecular, biological, immunological, and/or clinical aspects of the neurological complications of viral infections of the human nervous system which could include HIV-1. The program would conduct its work in conjunction with the Laboratory of Molecular Medicine and Neuroscience which has established areas of research on the infectious etiology of neurodegenerative diseases. The individual would have a demonstrated background and knowledge in virus-cell interactions using molecular technologies to investigate multifactorial mechanisms of damage to the brain including viral proteins, inflammatory molecules and chemokines. Experience with animal models of viral infections in the brain would be desirable. The candidate will have an earned Ph.D. and/or M.D. degree with excellent scientific skills in structuring an original and productive research program using outstanding communication and collaborative abilities. An individual selected for a tenure-track position is expected to build a dynamic and productive research group.

Candidates for a tenured position must have an international reputation and well-documented evidence of ongoing independent accomplishments. Laboratory facilities, start-up and sustained research funds and salary will be competitive with premier academic institutions.

Applicants should send curriculum vitae, bibliography, statement of research interests, and names of references to: **Dr. Story Landis, Director, NINDS, c/o Peggy Rollins, Office of the Scientific Director, Division of Intramural Research, NINDS, Building 35, Room GA908, NIH, Bethesda, MD 20892**. Or you can call 301-435-2232. The NINDS is one of the Institutes of the National Institutes of Health, a component of the Department of Health and Human Services. Applications will be reviewed upon receipt.



Director, Division of Prevention and Population Sciences National Heart, Lung, and Blood Institute

The National Heart Lung and Blood Institute (NHLBI) at the National Institutes of Health (NIH) seeks a dynamic scientist to provide strategic leadership for its newly reorganized Division of Prevention and Population Sciences (DPPS). The Director will assume primary responsibility for creating, nurturing and supporting internationally-renowned programs in population sciences and prevention in the areas of cardiovascular disease, and collaborating with closely aligned programs in the Division of Cardiovascular Diseases, Division of Lung Diseases, the Division of Blood Diseases and Resources, and the Division of Extramural Research Activities. The NHLBI also manages a strong program in biostatistics and resources for the conduct of clinical research. The NHLBI is engaged in a strategic planning process which will guide its scientific agenda for the next decade. The DPPS Director will advocate for areas of profound importance to the national and global populace, to establish and implement programs congruent with NHLBI's strategic plan and which will impact the health of the public. Functioning as a key member of the senior leadership team of the Institute, the incumbent will have a profound impact upon the national investment in research and the quality of service to the international research community. Applicants must possess an M.D., Ph.D. or equivalent degree as well as senior level research experience and ability to interact successfully with a broad range of individuals. The successful candidate will be a respected, accomplished scientist with maturity, integrity and outstanding communication skills.

Application Process: Please submit your CV, bibliography, and two letters of recommendation to: **Joanna Fesler, Program Manager, STG International, Inc, 4900 Seminary Rd., Suite 1100, Alexandria, VA 22311.** For further information, please call **877-784-6452** or email **jfesler@stginternational.com**. Your application package should be received by **October 15, 2006**. All information provided by candidates will remain strictly confidential and will not be released outside the NHLBI search process without a signed release from candidates.

Salary is commensurate with experience and a full package of Civil Service benefits is available including retirement, health and life insurance, leave and savings plan (401K equivalent).

The National Heart, Lung, and Blood Institute (NHLBI) provides leadership for a national program in diseases of the heart, blood vessels, lung, and blood; blood resources; and sleep disorders. With nationwide responsibility for improving the health and well-being of all Americans, the Department of Health and Human Services oversees the biomedical research programs of the NIH. The NIH encourages the application and nomination of qualified women, minorities and individuals with disabilities.



Director, Division of Cardiovascular Diseases National Heart, Lung, and Blood Institute

The National Heart Lung and Blood Institute (NHLBI) at the National Institutes of Health (NIH) seeks a dynamic physician-scientist to provide strategic leadership for its newly organized Division of Cardiovascular Diseases (DCVD). The Director will assume responsibility for creating and nurturing internationally-renowned programs which will participate actively in international research in cardiovascular diseases across the spectrum of basic science and clinical research including translational research and the conduct of a wide variety of clinical trials. The Director will recruit scientists and scientific administrators, develop and nurture a strong workforce, and build depth in disease-specific branches. Key challenges include establishment of priorities, integration of basic and clinical science, building teams, and interaction with scientific colleagues in many settings. Functioning as a key member of the senior leadership team of the Institute, the incumbent will collaborate with closely aligned programs in the Institute. The DCVD Director will have a profound impact upon the national investment in research, and the quality of service to the international research community. The Director of DCVD will have the opportunity to advocate for areas of critical importance to the national and global populace, to establish and implement programs congruent with NHLBI's strategic plan, and to improve the health of the public. Applicants must possess an MD or equivalent degree as well as senior level research experience, interpersonal and communications expertise and ability. The successful candidate will be a respected, accomplished researcher with maturity, integrity and outstanding communication skills.

Application Process: Please submit your CV, bibliography, and two letters of recommendation to: **Joanna Fesler, Program Manager, STG International, Inc, 4900 Seminary Rd., Suite 1100, Alexandria, VA 22311.** For further information, please call **877-784-6452** or email **jfesler@stginternational.com**. Your application package should be received by **October 15, 2006**. All information provided by candidates will remain strictly confidential and will not be released outside the NHLBI search process without a signed release from candidates.

Salary is commensurate with experience and a full package of Civil Service benefits is available including retirement, health and life insurance, leave and savings plan (401K equivalent).

The National Heart, Lung, and Blood Institute (NHLBI) provides leadership for a national program in diseases of the heart, blood vessels, lung, and blood; blood resources; and sleep disorders. With nationwide responsibility for improving the health and well-being of all Americans, the Department of Health and Human Services oversees the biomedical research programs of the NIH. The NIH encourages the application and nomination of qualified women, minorities and individuals with disabilities.



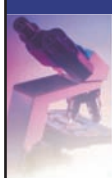
THE NATIONAL INSTITUTES OF HEALTH

OPPORTUNITIES @ NIH



THE NATIONAL INSTITUTES OF HEALTH

OPPORTUNITIES @ NIH



**Look into
INRO**
The Intramural NIAID
Research Opportunities Program

Help Today's Science Student Become Tomorrow's Top Researcher

The National Institute of Allergy and Infectious Diseases (NIAID), Division of Intramural Research (DIR), is seeking applicants for its Intramural NIAID Research Opportunities (INRO) program. INRO, currently in its fifth year, is a 5-day exploratory program intended for underrepresented minority students (i.e., American Indians or Alaska Natives, Native Hawaiians or other Pacific Islanders, blacks or African Americans, and Hispanics or Latinos) who are interested in a research career in the areas of allergy, immunology, and infectious diseases.

This 5-day program will be held February 4–8, 2007, at the National Institutes of Health in Bethesda, Maryland. Applicants to this highly competitive program should be undergraduate juniors or seniors interested in postbaccalaureate biomedical research, doctoral candidates seeking a postdoctoral training position, and first-year medical students contemplating a year-off research position.

Students selected to attend the INRO 2007 program will:

- **Learn about the Institute and about training opportunities available at NIAID** • **Listen to scientific lectures by world-renowned NIAID researchers** • **Tour Institute laboratories** • **Interview with principal investigators for potential training positions**

The INRO program will pay all expenses for travel, hotel accommodations, and meals. Eligible students include U.S. citizens or legal U.S. residents who belong to a minority population that is underrepresented in the sciences. Only students with strong academic standing will be considered.

For more details about student eligibility, highlights of last year's INRO program agenda, student testimonials from past programs, and detailed information about the application process, students may visit our Web site at www.niaid.nih.gov/labs/training/inro. Completed applications **MUST** be received by **October 15, 2006**.

For additional information, contact **Wendy J. Fibison, Ph.D., Associate Director, DIR, NIAID**, wfibison@niaid.nih.gov.



NIH *physician specialty training*

If your goal is to pursue academic medicine, and you have a U.S. medical license, consider fellowship or residency training with the physicians who are performing basic research and designing pivotal trials that will determine state-of-the-art care for the next decade. The National Institutes of Health is home to 250 research beds. We offer outstanding clinical training and a research experience that includes intense exposure to designing and analyzing clinical trials, and to cutting-edge basic science investigation, within our ACGME-accredited programs.

Student Loan Repayment of up to \$35,000 per year may be available to qualified individuals accepted into ACGME-accredited training programs.

For more information on ACGME-accredited programs and the Loan Repayment Program, visit our web sites at www.training.nih.gov, www.lrp.nih.gov

Office of Intramural Training and Education
Bethesda, Maryland 20892-1158
(888) 695-5343

NIH is dedicated to building a diverse community in its training and employment programs.

subspecialty training

Internal Medicine
Allergy and Immunology
Critical Care Medicine
Endocrinology and Metabolism
Hematology
Infectious Diseases
Medical Oncology
Rheumatology

Pediatrics
Allergy and Immunology
Endocrinology
Hematology/Oncology
Pediatrics/Medical Genetics

Blood Banking/Transfusion Medicine
Cytopathology
Hematopathology
Medical Genetics

residency training

Anatomic Pathology
Dermatology (*third year only*)
Psychiatry (*fourth year only*)

resident and medical student electives

Clinical and research electives for residents.
Flexible duration, 4 weeks, 8 weeks, up to one year.



THE NATIONAL INSTITUTES OF HEALTH

OPPORTUNITIES @ NIH

I want
to help
find a
cure for
cancer

Not worry
about my
student
loans

If you're interested in a biomedical research career, you should know that the National Institutes of Health Loan Repayment Programs may repay up to \$35,000 per year of your qualified educational loan debt.

Deadline for Applications is December 1

For more information, visit www.lrp.nih.gov or call 1-866-849-4047



The NIH Loan Repayment Programs



Healthier Futures Through Research

MOVERS

Sam Aronson, director, Brookhaven National Laboratory, Upton, New York



2006: Interim director, Brookhaven National Laboratory, Upton, New York

2005–06: Associate laboratory director for high-energy and nuclear physics, Brookhaven National Laboratory

2001–05: Chair, physics department, Brookhaven National Laboratory

Sam Aronson attributes his success in administration, in part, to a piece of fatherly advice. Even though Aronson was interested in science and engineering, his father suggested that a liberal arts undergraduate education would serve him well later in life. He took that advice, and went to Columbia College in Chicago, gaining a wide perspective of science's role in society, as well as a major in physics.

He then pursued a PhD in particle physics at Princeton University, detailing the properties of elementary particles called kaons. As a postdoc at the University of Chicago's Enrico Fermi Institute, he continued his kaon work, related to the observed imbalance between matter and antimatter.

Next, he accepted an assistant professorship at the University of Wisconsin, but opportunities for tenure in the physics department were limited. Aronson turned down offers from other universities to go to Brookhaven National Laboratory — his most pivotal career move.

He was lured there in the late 1970s by the prospect of building ISABELLE, a proton collider. After several years, the project was cancelled in favour of the superconducting supercollider, but this was also ultimately cancelled. Meanwhile, Aronson was helping build new experiments to measure the elastic scattering of neutrinos at Brookhaven, while also assisting with the Tevatron accelerator at Fermilab, which discovered the 'top' quark.

While deputy chair of the physics department at Brookhaven, Aronson helped design the Relativistic Heavy Ion Collider (RHIC), which uses heavy-ion collisions to study quarks and gluons at high temperatures. Aronson then moved to manage one of RHIC's detector construction projects, named PHENIX. "With RHIC, Brookhaven bounced back in a creative and spectacular way," he says.

After 30 years at Brookhaven, Aronson has taken the helm as director. "We were looking for the perfect combination of researcher and administrator," says Shirley Strum Kenny, Stony Brook University president and board chair of Brookhaven Science Associates, which operates the lab for the US Department of Energy.

Aronson will have to reposition Brookhaven's scientific agenda once RHIC faces competition from the Large Hadron Collider in Switzerland. "We face financial, political and sociological challenges to realize our vision of conducting basic-energy and nuclear physics," he says. "What I do now is not about hard science, but about managing science in a cultural and societal context." ■
Virginia Gewin

SCIENTISTS & SOCIETIES

Stand up and be quoted

An unconventional guide launched this month aims to help young researchers get involved in media coverage of science.

The guide, *Standing up for Science*, has been produced by a group of UK graduates in the Voice of Young Science (VoYS) network. We asked our fellow researchers what they thought about talking to the media. The answers were disappointing: no time, no opportunity, not enough experience, said many. Although a few respondents were more positive, most were sceptical. The public isn't interested in science, they said, and journalists only want to misquote us.

But at earlier VoYS workshops, science journalists from *The Times*, *The Guardian*, BBC Radio 4, and others voiced rebuttals to this. Yes, journalists need to sell newspapers, so they have to make science stories exciting. But by and large they want to report accurately. Their task is complicated by deadline pressures, and little or no control over headlines.

Scientist-media interaction would benefit from enthusiastic early-career researchers eager to make science exciting and relevant to the public. Anna Fazackerley, who has written for *The Guardian* and the *Financial Times*, suggested that young people could bring a fresh perspective articulated

in informal, engaging language.

In *Standing up for Science*, we tried to help researchers understand the media's agenda and objectives, even if they feel they're not cut out to be communicators themselves. We conducted interviews with newspaper and radio journalists, press officers and scientists who had talked to the media. We put the most interesting conversations into an informal guide to address our peers' concerns.

Some worried that journalists would 'sex up' their stories. "I was horrified when they asked, 'Can we say you've discovered anti-gravity?'," said Emma King, a cosmologist at the University of Nottingham. Scientists must be prepared to deal with hype and take control in an interview, advised scientists and press officers.

On the other hand, countered Mark Henderson of *The Times*, "If we overplay something it is usually because somebody has overplayed it to us."

Scientists should become media-savvy, asking journalists about their deadline, their audience and the angle of the story. The more such scientists there are, the more accurate and interesting science coverage will be. ■

Richard Van Noorden is a member of the VoYS network and a former intern at Nature.

GRADUATE JOURNAL

D-day

After barricading myself in my office every day and night for the past couple of months, I've finally completed a first draft of my dissertation. A dissertation defence date has been set, and friends and family have made arrangements to travel to Hawaii to support me as I attempt to survive the final battle of my graduate career. The sun and surf of Hawaii are probably the real reasons many of them are visiting. I doubt if they want hear about the biogeochemical consequences of rising CO₂ and 'ocean acidification' in carbonate ecosystems. But my parents would have made it regardless — whether I lived on the North Pole or Hawaii, whether I was presenting the first evidence of life on Mars or just reading out loud from a phonebook.

As my defence day rapidly approaches, I still have one big problem: how do I fit five years of research into a 45-minute talk? What should I include and what should I exclude? My self-inflated science ego tells me to include as much as I possibly can. After all, I've devoted considerable effort to my dissertation and this is my chance to demonstrate what I've accomplished. Furthermore, it hurts tremendously every time I cut something out of my presentation. I'm hopeful that I'll overcome this problem and present a memorable story rather than an incoherent cluster of results. ■

Andreas Andersson is a final-year PhD student in oceanography at the University of Hawaii.

Operation Tesla

Try to blend in.

Jeff Hecht

An hour ago, Frankel had been on the holovision stage, being introduced as part of Team Beta.

"It is 200 years since the birth of Nikola Tesla in 1856," the announcer had said. "Now Enigmas of the Twentieth Century is sending three teams of intrepid time travellers to find the legendary inventor's lost papers on wireless transmission, death rays and energy before they went missing."

Now Frankel was back in Tesla's Manhattan, walking north on Sixth Avenue from 40th Street towards the entrance to Bryant Park. Massive yellow taxis zoomed by. A blue-uniformed policeman checked him out, then walked on past, but Frankel still worried. "I haven't heard from Watkins," he said to the phone transmitter buried in his collar.

He could hear Johnson's annoyance through his earpiece. "I told you dead spots are inevitable, even with the main transceiver in the Empire State Building. I know these phones are outdated for 2056, but they're hidden so nobody can see them. Everything you're wearing belongs in October 1937."

"Okay," Frankel agreed, reluctantly. He was grateful the rules let them use technology from any time in the twentieth century. It would have been hard to hide mobile phones built with vacuum tubes. Getting the colours of women's fashions would have been hard, but the three men on Team Beta could get by with dark suits, ties and white shirts. Yet that hadn't helped all-male Team Alpha, who hadn't come back on time.

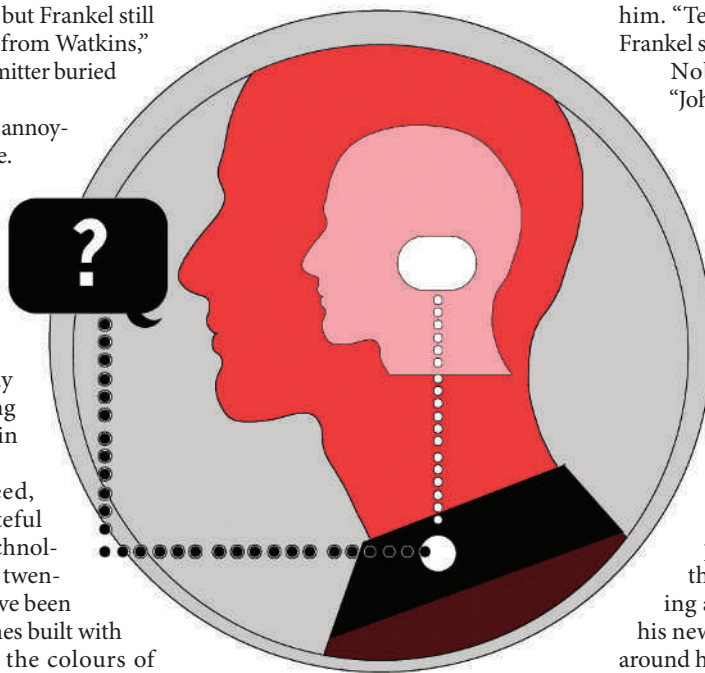
"Do you see Tesla yet?" Johnson's voice hissed through the earpiece. "He should be near the public library, feeding the pigeons."

Inside the park, Frankel looked east toward the library. A tall, gaunt elderly man on a bench scattered crumbs to a flock of head-bobbing pigeons. Frankel recognized the inventor. He wondered what secrets of power-beaming and wireless transmission were hidden in the missing papers. "I see him. Where's Watkins? I haven't heard from him in 15 minutes."

"That's when he went into the Hotel New Yorker. I watched him from Cut Rate Drugs across Eighth Avenue; I'm having a soda inside. The hotel is a complex 43-storey building, and something inside must be blocking his signal. I'll tell him you've spotted Tesla. If I can't reach him on his mobile, I'll head up to Tesla's room myself. I have a set of tools to get in."

"But what if I can't reach you?" Frankel worried. The operation depended on everyone keeping in touch so they could copy the papers and bring them to the retrieval point. They'd lose points if they had to be retrieved from different locations.

"Don't worry. You're perfectly safe. Nobody can eavesdrop on these calls. Just watch Tesla."



"OK," Frankel agreed. Watching him was the easiest part of the operation. The goal of Operation Tesla was to bring back copies of Tesla's missing papers: Enigmas of the Twentieth Century promised fame and a handsome reward for the team that came back with the best. Watkins and Johnson thought the documents were in the hotel room where he lived. The old-fashioned mechanical locks should be easy to pick; the challenge was to find and copy the right papers.

Frankel settled down on a bench to keep a discreet eye on Tesla. The inventor was engrossed in his pigeons, as if he knew each one. Frankel opened a copy of the *New York Sun* he had bought from a news

vendor. The headlines ranged from world politics to sports. Roosevelt had talked on the transatlantic radio-telephone with the British Prime Minister Neville Chamberlain about Adolf Hitler's meeting with Benito Mussolini. Police were questioning the staff of Bellevue Hospital about three lunatics who had vanished without trace from a locked ward. The Yankees had just beat the New York Giants four games to one in the World Series, and a sports writer wondered how long Lou Gehrig could keep playing.

Glancing at his authentic 1930s watch, Frankel saw it was time for another report. He scanned the park, and noted the inventor feeding the pigeons and three policemen walking down the path that went by him. "Tesla is still feeding the pigeons," Frankel said to his collar.

Nobody replied. That was odd. "Johnson! Watkins! Are you there?"

Worried that the police might have caught Watkins and Johnson trying to break into Tesla's hotel room, Frankel tried to look inconspicuous behind his paper. "Johnson, Watkins? Where are you? Rendezvous is in one hour and 15 minutes!"

A heavy hand clamped down on Frankel's right shoulder, and as he started to jump, a second clamped on his left. "All right, pal, we'll take you back to a nice, quiet padded cell," said one of the three policemen suddenly standing around him. They pulled away his newspaper and snapped handcuffs around his wrists, then patted him down, pulling out his wallet but not spotting the little transmitter sewn into his jacket.

"Johnson, Watkins, help!" He cried into his collar.

The first policeman shook his head. "If those are your pals from Bellevue, they're already at the precinct, talking to themselves — just like you."

Frankel shivered. "They caught Team Alpha, and now they've got me," he said into his collar phone. "I don't know what tipped them off."

The policeman rolled his eyes. ■
Jeff Hecht writes regularly for *New Scientist* and *Laser Focus World*. His most recent book is *Beam: The Race to Make the Laser*. He still uses a landline phone.

JACEY